第 39 卷　第 2 期　天 文 学 进 展　Vol. 39,　No. 2
2021 年 6 月　PROGRESS IN ASTRONOMY　June., 2021

doi: 10.3969/j.issn.1000-8349.2021.02.02

# Stäckel 引力势理论的拓展及其应用

顾建宇[1,2]，董小波[1,2]

(1. 中国科学院 云南天文台，昆明 650216； 2. 中国科学院大学，北京 100049)

**摘要：** Stäckel 引力势是一类最普遍形式的可分离势。具有 Stäckel 形式的星系是完全可积系统，即其中的恒星轨道都是规则的，其运动积分可以解析求出。运动积分——特别是作用量 ($\boldsymbol{J}$，一种特殊的运动积分)，能简化恒星运动的描述；研究星系中恒星的运动学和动力学的重要途径是作用量空间，比如使用以作用量作为参量的分布函数 $f(J_R, L_z, J_z)$。通过将一般星系的引力势局部地近似成 Stäckel 势，人们可以估算一般星系中恒星运动的作用量。介绍天文学家在拓展 Stäckel 引力势的研究上取得的重大进展：在基于 Stäckel 势理论估算一般星系中的作用量 (或运动积分) 方面，开发出了若干快速数值算法 (例如 Stäckel 作弊法 [Stäckel Fudge]、Stäckel 拟合法)，基于环面映射的收敛性高精度数值算法 (例如轨道积分拟合生成函数法、迭代环面构造法)，以及提出作为近似的运动积的解析表达式；在分布函数的构造方面，基于上述估算的作用量或直接利用 Stäckel 势的运动积分公式 (主要是 $I_3$ 表达式)，提出了一些分布函数模型，$f(J_R, L_z, J_z)$ 或 $f(E, L_z, I_3)$。这些进展使得人们可以对星系开展基于分布函数的建模。此外，还介绍近几年把这些方法应用到银河系观测数据的若干代表性工作。

**关　键　词：** Stäckel 引力势；作用量-角变量估算；银河系分布函数建模

**中图分类号：** P135　　**文献标识码：** A

## 1　引言和理论基础

银河系研究的主要目标之一，是测量并计算出银河系的结构，即各种组成成分 (如恒星核球、恒星盘、暗物质晕和气体等) 的空间分布，以及它们的速度空间信息。为了实现这个目标，在观测上，需要高精度的天体测量学及测光学数据 (如 Gaia) 和光谱数据 (如 LAMOST)；在理论和计算上，星系动力学建模是核心。星系动力学建模的理论基础是恒星系统的动力学，即研究大量的点质量粒子在它们共同的引力场中的运动[8]。关于银河系建模，尽管星系作为自引力系统，没有星系能处于严格的热力学平衡态或完美的动力学平衡态，然而平衡态动力学模型 (或称稳态 [steady state]，或准定态 [quasi-stationary state])，

**收稿日期：** 2020-07-06；　**修回日期：** 2021-02-28

**资助项目：** 国家自然科学基金 (11873083, 11773074)

**通讯作者：** 董小波，xbdong@ynao.ac.cn

对于解释银河系和河外星系的观测至关重要。从实际情况看，暗物质晕作为星系引力势的主要提供者，仅在短暂且罕见的主并合事件中，才远离动态平衡，一般情况下该动态平衡仅受吸积或次并合等事件的小扰动。从模型需求的角度看，由动力学测量数据推断出物质分布，该模型要求假设星系处于 (准) 平衡态；另外，平衡态模型是最简单的模型，更复杂的情况可以被建模为平衡态模型的“扰动”、“变形”或迭加[20]。

运动积分是粒子在轨道运动中的不变量；作用量，则是可以与角变量形成正则坐标的一类特殊的运动积分。由于在“作用量-角变量”坐标下运动方程形式的简单性，作用量成为很有吸引力的变量。并且作用量无论在微扰论方法上还是分布函数的建模上，都是十分必要的。但是，粒子在一般引力势 (例如真实的星系) 中的运动，基本上不能得到解析的运动积分或作用量；仅当引力势为可分离势时，能解析地得到。Stäckel 势是椭球坐标系下可分离的引力势形式，该引力势下的运动积分可被解析地计算[1]。在应用上，通过将一般引力势近似或拟合为 Stäckel 势，人们可估算一般引力势下的运动积分。近十几年来，一些研究者基于Stäckel引力势理论中的解析式，用 (参见 3.2 节) 解析或数值的 (例如 Binney 的工作) 方法来作拓展，即基于 Stäckel 势来估算一般星系中的作用量 (或运动积分)。特别是最近几年来，Binney 和他的合作者们把他们的方法都实现为计算机程序，并以公开源码的方式释放了 (详见第 2 章)。

Stäckel 势作为一个简单的引力势形式，早在 19 世纪[9, 10] 就已被给出。在星系动力学的理论推导方面，最普遍形式的三轴 Stäckel 势在 30 多年前就已经得到了详尽的总结，包括三个孤立运动积分 (或作用量) 的解析表达式，恒星轨道的分类等[1]。但直到近几年，有关 Stäckel 势的知识才在实际的银河系建模中应用起来。这主要归功于上段所述的 Stäckel 势拓展工作，即估算一般引力势的作用量 (或运动积分) 的方法和程序[15, 54]。

接下来，第 1 章的其他小节分别介绍：恒星系统的动力学及统计力学的概要、运动积分以及 Stäckel 势的相关基本背景知识[1]。第 2 章介绍估算一般引力势的运动积分和作用量的方法：2.1 节介绍 Stäckel Fudge 作弊法[13, 20, 22, 24]；2.2 节介绍 Stäckel 引力势拟合法[19]，以及更早期的工作；2.3 节介绍一类环面坐标变换的作用量估算方法，基于 Stäckel 玩具势模型得到一般引力势下的作用量[20, 21, 24]；2.4 节是我们基于文献 [24] 所介绍的 TACT 程序对各种方法做的比较；2.5 节则是介绍一些其他求运动积分的方法和程序。第 3 章介绍近几年来这些方法在银河系数据上的应用 (如文献 [44])。第 4 章进行讨论和总结，特别是简介“作用量-角变量”方法在不可积的星系现象 (见 4.2 节) 和非平衡的星系过程 (见 4.3 节) 中应用的最新进展。

### 1.1 恒星系统的理论基础

本文讨论的单粒子在整体势的运动，是从引力多体运动简化而来的，即单粒子近似或称平均场近似[60]。三维背景空间中一个包含有 $N$ 个粒子的经典系统，可用哈密顿方程描述：

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \ , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \ , \quad i = 1, 2, 3, ..., 3N \ . \tag{1}$$

对于该系统所有粒子的坐标和动量所构成的全 $6N$ 维相空间 (一般称为 $\Gamma$ 相空间)，通过

相空间代表点个数的局部守恒可推导得，其概率密度 $\rho(q,p)$ 满足刘维尔方程：

$$\frac{\partial\rho}{\partial t}+\{\rho,H\}_{\rm PB}=0 \quad , \tag{2}$$

其中，$\rho=\rho(q_i,p_i,t)$，$i=1,2,3,\cdots,3N$ 是系综意义的概率分布；$\rho$, $H$ 的泊松括号 $\{\rho,H\}_{\rm PB}=\sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial q_i}\frac{\partial H}{\partial p_i}-\frac{\partial H}{\partial q_i}\frac{\partial\rho}{\partial p_i}\right)$。式 (2) 中，左边其实就是拉格朗日视角的随流导数 $\mathrm{d}\rho/\mathrm{d}t$，即其相体积不变；这就是刘维尔保体积定理。

下面用 $\boldsymbol{x}=x_j,\boldsymbol{p}=p_j\ (j=1,2,3)$ 代表单个粒子的坐标和动量；$\boldsymbol{q}_k,\boldsymbol{p}_k$ 代表第 $k$ 个粒子的坐标和速度；$f_n$ 代表 $n$ 粒子联合分布函数。对上述刘维尔方程作单粒子近似 $f_2(\boldsymbol{q}_1,\boldsymbol{p}_1,\boldsymbol{q}_2,\boldsymbol{p}_2,t)=f_1(\boldsymbol{q}_1,\boldsymbol{p}_1,t)f_1(\boldsymbol{q}_2,\boldsymbol{p}_2,t)$[①]，可得到玻尔兹曼方程[②]：

$$\frac{\partial f}{\partial t}+\boldsymbol{v}\cdot\nabla_q f+\boldsymbol{F}\cdot\nabla_p f=\int\mathrm{d}\boldsymbol{p_i}\int\mathrm{d}\Omega\quad\sigma(\Omega)\|\boldsymbol{v}-\boldsymbol{v_1}\|(f'f_1'-ff_1) \quad , \tag{3}$$

其中，这里 $f=f(\boldsymbol{q},\boldsymbol{p},t)$ 是单粒子分布函数，即处于自由度 3 的坐标和速度所构成的 6 维相空间 (一般称为 $\mu$ 相空间)；$\boldsymbol{v}$ 代表速度，$\boldsymbol{F}$ 代表受力[③]；$\sigma(\Omega)$ 是分子微分散射截面，这里考虑了二体碰撞；对于其他作用类型 (如刚性) 的散射，将可能不再是这样的形式。玻尔兹曼方程体系中所需要的假设非常繁杂 (甚至冗余)，而且至今仍有难以严格数学证明的规律。本文对玻尔兹曼方程仅作介绍而不作分析 (更多参见文献 [8] 中的第 4, 7 章和文献 [60] 中的第 3, 4 章)。此外可以验证，刘维尔方程在 $\varGamma$ 相空间的分布函数的熵是不变的，而玻尔兹曼方程在 $\mu$ 相空间的分布函数的熵是不减的。

对玻尔兹曼方程进行局域平衡假设，碰撞项将消失。其实，在星系这种具体的系统中，恒星的半径与星系尺度相比很小，其平均自由程远大于系统尺度，碰撞时间远大于穿越整个星系的时间，比宇宙年龄还大很多量级，因此碰撞项本应忽略不计。把玻尔兹曼方程两边乘上一个碰撞不变量 $\xi(\boldsymbol{x},\boldsymbol{v},t)$ (如质量、动量、能量等)，然后作局域平均 (由于碰撞前后的某些对称性, 局域平均后碰撞项不做贡献，见文献 [60] 第 7A 章)，可分别得到流体力学性质的方程 (注：流体力学过程是局域平衡过程)。例如对于动量守恒，玻尔兹曼方程两边乘以 $\xi(\boldsymbol{x},\boldsymbol{v},t)=mv_i(i=x,y,z)$，通过对动量空间积分并利用动量空间的散度定理 (无穷远处分布函数为 0，$\lim_{\boldsymbol{p}\to\infty}f=0$)，可得到

$$\frac{\partial}{\partial t}\langle\rho_i v_i\rangle+\nabla_x\cdot\langle\rho v_i\boldsymbol{v}\rangle-\frac{1}{m}\rho F_i=0 \quad . \tag{4}$$

此式实质上等同于流体力学动量守恒方程：

$$\frac{\partial(\rho\boldsymbol{u})}{\partial t}+\nabla_x\cdot(\rho\boldsymbol{u}\boldsymbol{u}+\underline{P})=\frac{\rho}{m}\boldsymbol{F} \quad . \tag{5}$$

---

①这是一种平均场近似，即将粒子间相互作用当成总背景场，而不考虑粒子间相互作用和关联；体现在分布函数上就是每个粒子的单粒子分布是无关联的，因此乘积为它们的联合分布。

②由玻尔兹曼于 1872 年提出。从刘维尔到玻尔兹曼方程的约化，中间还有一个 BBGKY 级联方程。

③该受力是对于该“单粒子”而言的，施力者为“其他粒子”(实际上是分布函数) 的作用 (因而在这里直接体现平均场近似的思想)，以及可能存在的外场。玻尔兹曼方程可以用来描述非自洽的体系 (自洽体系是指粒子所受力全部来自于研究系统本身，如自引力系统)，即通过建立多种组分的玻尔兹曼方程来描述。

其中，$\underline{P}$ 为压力张量。流体力学的质量、能量守恒也可以类似地得到 (见文献 [60] 第 7 章)。

这样，我们就有了“$N$ 体系统 (力学描述)-刘维尔方程 ($6N$ 维)-玻尔兹曼方程 (6 维)-流体力学方程 (流场)”的联系。

另一方面，从无碰撞玻尔兹曼方程出发，两边乘以速度分量，再对速度空间积分并利用散度定理 (与得到式 (4) 的方法类似)，我们得到 Jeans 方程：

$$\nu\frac{\partial\bar{v}_j}{\partial t}+\sum_i\nu\bar{v}_i\frac{\partial\bar{v}_j}{\partial x_i}=-\nu\frac{\partial\Phi}{\partial x_j}-\sum_i\frac{\partial(\nu\sigma_{ij}^2)}{\partial x_i}\quad. \tag{6}$$

其中，$\nu$ 代表位形空间数密度，$\sigma_{ij}$ 为速度弥散张量。Jeans 方程不仅通过平均简化了方程，还将分布函数与这些可观测量联系起来，从而对联系观测天文的工作很有帮助。Jeans 方程还给人们提供了一种检测分布函数正确性的方法。当然, Jeans 方程仍然不是闭合的。

在 $\mu$ 相空间中的粒子分布函数 $f=f(\boldsymbol{x},\boldsymbol{v},t)$ ①，跟粒子视角的单粒子近似下的运动 ($\boldsymbol{x}(t)$, $\boldsymbol{v}(t)$) 有着密切的联系。这种单粒子分布函数，实际上假设体系是由无穷多的点粒子构成，某种程度上与场论的意义比较像 (场论就是处理无穷多自由度)。从场的角度来理解，这是把所有粒子之间的相互作用当做背景场，然后对于任一粒子，只须研究它在这个场中的运动，不必考虑其他粒子对它的作用；此即前述平均场的思想。但是，实际上想详细了解星系结构与演化 (如分布函数)，直接从自引力系统的统计物理方程出发去求解分布函数并不容易，而且运用最大熵原理之类的方法也是不够的 (由于星系作为长程相互作用系统的特性，见第 4 章)。为得到分布函数，一种方法是 $N$ 体模拟，基于模拟的结果作统计。另一种方法是根据粒子在引力势的运动特性来构建分布函数模型[8]，以及其他理论手段 (详见文献 [8])；在这种方法中，对运动积分的研究是个基础的环节。下面，我们主要从哈密顿力学以及单粒子近似下的恒星轨道运动的角度出发，介绍恒星在静态引力势下的运动积分的相关内容。

## 1.2 运动积分和作用量-角变量体系简介

### 1.2.1 哈密顿-雅克比方程

力学系统在相空间的正则坐标 ($\boldsymbol{q}$, $\boldsymbol{p}$) 的演化，可以看做是无穷多个连续以哈密顿量为生成函数的无穷小正则变换的结果。在正则变换中令新的哈密顿量为 0，并选取第二类生成函数 $F_2$②可使某种新的正则坐标 $\boldsymbol{Q}$, $\boldsymbol{P}$ 均为常数，即 $\boldsymbol{Q}=\boldsymbol{\alpha}$, $\boldsymbol{P}=\boldsymbol{\beta}$，可得正则变换的方程：

$$H\left(\boldsymbol{q},\frac{\partial S}{\partial\boldsymbol{q}},t\right)+\frac{\partial S}{\partial t}=0\quad, \tag{7}$$

被称为哈密顿-雅克比方程 (哈雅方程)，其解的形式为 $S=S(\boldsymbol{q},\boldsymbol{\beta})$ 就是生成函数 $F_2$。另外，

---

① 由于 $\boldsymbol{p}=m\boldsymbol{v}$ ($m$ 为单粒子的质量)，而且在星系动力学的粗糙建模中常常将测试天体考虑为无穷多同质粒子的分布，且为方便一般假设测试粒子的质量为 1，于是下面我们这里的习惯，以下讨论的 $f$ 都是关于 $\boldsymbol{x},\boldsymbol{v}$ 的分布函数。

② 第二类生成函数是指：在正则变换中，取独立的广义坐标变量为老的坐标和新的动量，亦即生成函数为老坐标、新动量以及时间的函数。

第二类正则变换中哈密顿函数之间的微分关系表示为：

$$p_i = \frac{\partial S}{\partial q_i} \ , \quad \alpha_i = \frac{\partial S}{\partial \beta_i} \ , \quad i = 1, 2, 3, \cdots, n \ . \tag{8}$$

1.2.2 运动积分

如果一个力学量不显含时间而且保持不变，$h(\boldsymbol{q}, \boldsymbol{p}) = \text{const}$，就称之为系统的运动积分，即运动积分是粒子在轨道运动中的不变量。寻找系统的这种不变量，具有重要意义。对于一个自治哈密顿系统，哈密顿量不显含时间。此时哈密顿量本身就是一个运动积分，其积分常数 $E$ 等于系统总能量。利用公式 (7) 和 (8)，哈雅方程的解可写成时间项和广义坐标项分离的形式，表示为：

$$S(\boldsymbol{q}, \boldsymbol{\beta}, t) = W(\boldsymbol{q}, \boldsymbol{\beta}) - h\,t \ , \tag{9}$$

式中的 $W$ 称为哈密顿特征函数。对于自由度 $n$ 的自治哈密顿系统，最多有 $n$ 个两两独立且对合[①]的运动积分，因此在上式的积分常数 $\boldsymbol{\beta} = \{\beta_1, ..., \beta_n\}$ 和 $h$ 这 $(n+1)$ 个常数中，至少有一个不独立，即通过 $h = h(\beta_1, ..., \beta_n)$ 相联系，我们取 $\beta_n = h$。对式 (9) 两边微分得到两个微分关系，以及生成函数的微分关系，即式 (8)；再将 $W$ 代入 $S$ 中，可得到关系 $W$ 与新旧相空间坐标的微分关系：

$$p_i = \frac{\partial W}{\partial q_i} \ , \quad \alpha_i = \frac{\partial W}{\partial \beta_i} \ , \quad i = 1, 2, 3, \cdots, n \ . \tag{10}$$

现在我们以哈密顿特征函数 $W$ 作为生成函数作正则变换 $(\boldsymbol{q}, \boldsymbol{p}) \to (\boldsymbol{Q}, \boldsymbol{P})$，其中，

$$\begin{cases} Q_i = \dfrac{\partial W(\boldsymbol{q}, \boldsymbol{\beta})}{\partial P_i} = \dfrac{\partial W(\boldsymbol{q} \ , \boldsymbol{\beta})}{\partial \beta_i} \\ P_i = \beta_i \ , \quad i = 1, 2, 3, \cdots, n \end{cases} . \tag{11}$$

此时的新哈密顿量仍是 $K = H$，并且新哈密顿函数只含新广义动量而不含新广义坐标，该正则变换是作为后面讨论运动积分坐标的基础。

更进一步，对于自治 (完全) 可分离的生成函数 $W$，可写为：

$$W = \sum_i W_i(q_i, \boldsymbol{\beta}) \ , \tag{12}$$

则相应地可以分离出 $n$ 个哈雅方程 (不对 $i$ 求和)，可表示为：

$$H_i\left(q_i, \frac{\partial W_i}{\partial q_i}, \boldsymbol{\beta}\right) = \beta_i \ . \tag{13}$$

哈雅方程方法正可方便可分离系统的求解。

如上所述，若自由度 $n$ 的自治哈密顿系统为 (完全) 可积的，则它存在 $n$ 个互相对合的运动积分 $I_i (i = 1, 2, 3, \cdots, n)$ (我们用 $I_i$ 代表已经从 $E, \boldsymbol{\beta}$ 中选好的运动积分)；此即 (刘维

[①]如果两个运动积分的泊松括号为 0，那么它们是对合的。

尔) 可积性定理的内容。另一方面，根据刘维尔保体积定理，以这些运动积分为某一组常数时的坐标集合，$M = \{(\boldsymbol{q},\boldsymbol{p})|I_i(\boldsymbol{q},\boldsymbol{p}) = \beta_i\}$，它在以 $I_i$ 为哈密顿函数的相流下保持不变，并且同胚于 $n$ 维环面 $T^n$ (若不存在共振)，因此将 (完全) 可积系统的轨道转化为一个 $n$ 维准周期运动[63]。

关于运动积分，值得注意两点：第一，在星系动力学语境中，运动积分和运动常量是作了严格区分的两个概念；第二，关于运动积分还有更多细分的概念。对于恒星在星系引力势中的运动，给定引力场，运动常量的定义是：关于相空间坐标和时间的、沿恒星轨道不变的任意函数，即 $C(\boldsymbol{x}(t),\boldsymbol{v}(t),t) = \mathrm{const}$。运动积分也是在恒星轨道中保持常数的量 (即运动常量)，但它进一步要求是不显含时间的函数，$I(\boldsymbol{x}(t),\boldsymbol{v}(t)) = \mathrm{const}$。运动积分可以分为孤立积分和非孤立积分两类[8]。孤立积分能影响恒星轨道在相空间的分布，即相空间轨迹 (phase trajectory) 所占据空间的维度；对于可积系统，每增加一个孤立积分，就把相空间轨迹所占据流形 (即相空间的子空间) 的维度降低一维。非孤立积分则对于恒星轨道在相空间的分布没有影响，因而没有实用价值。

另外，在完全可积的或部分可积的系统中，请注意对合积分与孤立积分这两个概念的区别。对于自由度为 $n$ 的自治哈密顿系统，相空间为 $2n$ 维，最多有 $2n-1$ 个独立的 (即线性无关的) 运动积分，其中孤立运动积分最多可以是 $2n-1$ 个；而对合是与可积性有关的概念，两两对合的运动积分不超过 $n$ 个 (当对合积分的个数为 $n$ 时，该系统就是刘维尔当年所定义的完全可积的)。因此，若有一组运动积分是互相对合的，那么它们肯定是孤立积分；但反之不然。例如，对于球对称引力势中的恒星轨道 (6 维相空间)，运动积分有 5 个，其中孤立积分可以是 4 个或者 5 个 (仅当轨道为开普勒型时)，但相互对合的积分的个数是 3[8]。对于星系动力学中常见的轴对称引力势系统 (例如盘星系)，一般有 3 个孤立运动积分，能量 $E$、角动量的垂向 ($z$) 分量 $L_z$、以及第三个运动积分 $I_3$，但 $I_3$ 往往是没有解析形式的；对于三轴引力势系统，一般有 3 个孤立运动积分，能量 $E$ 和 $I_2, I_3$，但 $I_2, I_3$ 往往是没有解析形式的。过去，前人有很多工作就是在找第 2、第 3 运动积分，以及在研究稳态系统的恒星分布时所用的、以此为自变量的分布函数 $f(\boldsymbol{I})$ (这也是本综述文章的中心内容)。

#### 1.2.3　作用量-角变量体系

尽管运动积分 $\boldsymbol{I}$ 与其广义坐标 (记为 $\boldsymbol{\theta}$，角变量) 已经将运动化为准周期运动，但是运动积分一般不是正则坐标，即它与其广义坐标不满足哈密顿共轭，运动积分经过变换后能够满足共轭的叫做作用量 $\boldsymbol{J}$。下面我们对角变量和作用量的要求作一下简介 (可见文献 [8] 第3.5.1 节)。

如前所述，我们假定存在从 $\boldsymbol{I}$ 到 $\boldsymbol{J}$ 的可逆的变换 ($\det(\partial\boldsymbol{J}/\partial\boldsymbol{I}) \neq 0$)，使作用量 $\boldsymbol{J}$ 满足：(1) 运动积分，(2) 正则坐标。于是哈密顿方程 (1) 可用，关于 $\boldsymbol{J}$ 的部分有：

$$\dot{J}_i = -\frac{\partial H}{\partial \theta_i} = 0 \ . \tag{14}$$

于是哈密顿量只是作用量的函数，表示为：

$$\dot{\theta}_i = \frac{\partial H}{\partial J_i} \equiv \Omega_i(\boldsymbol{J}) \ , \quad \mathrm{const} \ \Rightarrow \theta_i = \theta_i(0) + \Omega_i(t) \ . \tag{15}$$

其中，$\Omega_i(\boldsymbol{J})$ 是角频率。对于束缚轨道，角变量 $\boldsymbol{\theta}$ 和运动积分所描述的就是一个 $n$ 维准周期运动，即前述的同胚于 $n$ 维环面 $T^n$ 的流形 $M$；$M$ 上某一维的任一闭合环路 $\gamma_i$ (同胚于圆周 $S^1$)，由某一组 $(\theta_i, J_i)$ 所代表，并且要求 $\theta$ 沿 $\gamma_i$ 的环路积分为 $2\pi$。

于是我们可以将相空间轨道展开成傅里叶级数：

$$\boldsymbol{x}(\boldsymbol{\theta}, \boldsymbol{J}) = \sum_{\boldsymbol{n}} X_{\boldsymbol{n}}(\boldsymbol{J})\mathrm{e}^{\mathrm{i}\boldsymbol{n}\cdot\boldsymbol{\theta}} = \sum_{\boldsymbol{n}} X'_{\boldsymbol{n}}(\boldsymbol{J})\mathrm{e}^{\mathrm{i}\boldsymbol{n}\cdot\boldsymbol{\Omega}t} \ , \tag{16}$$

角频率之间的比例关系丰富了体系周期类型和遍历性的复杂性 (见第 4 章)。天体动力系统的周期特性一般来说非常强，这也是作用量-角变量体系在天体力学中经久不衰的重要原因。

定义作用量 $J'_i$ 为 $\gamma_i$ 所围区域的相体积[8]，则有：

$$\begin{aligned} J'_i =& \frac{1}{2\pi}\int\int \mathrm{d}\boldsymbol{q}\cdot\mathrm{d}\boldsymbol{p} = \frac{1}{2\pi}\int\int_{\text{interior of }\gamma_i} \mathrm{d}\boldsymbol{\theta}\cdot\mathrm{d}\boldsymbol{J}v = \frac{1}{2\pi}\int\int_{\text{interior of }\gamma_i} \mathrm{d}\theta_i\mathrm{d}J_i = \\ & \frac{1}{2\pi}\left(\oint_{\gamma_i}\mathrm{d}\theta_i\mathrm{d}J_i - \oint_{J_i=J_i^c}\mathrm{d}\theta_i\mathrm{d}J_i\right) = J_i - J_i^c \ , \end{aligned} \tag{17}$$

其中第 2 个等式是由于正则变换的相体积不变性；第 4 个等式是转化为边界积分，并且刨除积分奇点的邻域。将奇点置于作用量-角变量坐标的中心，$J_i^c = 0$，此时 $J'_i$ 等于作用量的不变量。于是，我们由正则变换的庞加莱不变量的不变性，可得：

$$\boldsymbol{J} = \frac{1}{2\pi}\oint_{\gamma} J_i\mathrm{d}\theta_i = \frac{1}{2\pi}\oint_{T\gamma} \boldsymbol{p}\cdot\mathrm{d}\boldsymbol{q} \ , \tag{18}$$

其中，$T$ 是正则变换的矩阵，指在 $\gamma_i$ 对应的相空间所在区域。也由此可见，某个作用量的值 $J_i$ 代表子环 $\gamma_i$ 所围的面积 (当对 $q_i, p_i$ 选取合适的标度)，而 $\theta_i$ 代表转过的角度。

实际上，由式 (12) 以及式 (11) 的第 1 式，易得 $p_i = \dfrac{\partial W_i(q_i, \boldsymbol{\beta})}{\partial q_i}$，即对于完全可分离系统，$p_i$ 只依赖于相应的 $q_i$ 而独立于其他广义坐标。这样一来，作用量的计算可简化为 $J = \dfrac{1}{2\pi}\oint p_i\mathrm{d}q_i$。并且 $W_i$ 沿广义坐标 $\theta_i$ 运动一周而其他不变，回到初始状态，此过程中哈密顿特征函数 $W_i$ 的变化量 $\Delta W_i = \oint p_i\mathrm{d}q_i$，正好是作用量的 $2\pi$ 倍，这就是作用量的意义。

此外，我们可用式 (15) 计算角频率，用式 (11) 中第 1 行的哈密顿特征函数计算角变量的值。数值计算上见附录 A。

于是，从一般的相空间坐标到作用量的关系为：相空间坐标 $(\boldsymbol{q}, \boldsymbol{p}) \to$ 运动积分 $\boldsymbol{I}$ 和角坐标 $\boldsymbol{\theta} \to$ 正则环面坐标 $(\boldsymbol{J}, \boldsymbol{\theta})$。

#### 1.2.4 基于守恒量的分布函数

关于单粒子分布函数 $f(\boldsymbol{x}, \boldsymbol{v})$ (此时为 3 个自由度的，由坐标、速度所构成的 6 维 $\mu$ 相空间)，对于 (准) 稳态体系 (指 $\partial f/\partial t = 0$)，由无碰撞玻尔兹曼方程可推导，分布函数可写成运动积分的函数 $f = f(\boldsymbol{I})$，此即 Jeans 定理；而对于规则轨道具有非谐振频率的准稳态恒星系统的分布函数，可写成 3 个作用量的函数 $f = f(\boldsymbol{J})$，此即强 Jeans 定理。因此人们能更加不冗余、深刻地理解恒星的分布。

早先前人所用的分布函数模型，常常是 $f(E,\cdots)$ 的函数形式，即把能量作为分布函数的参数 (参见 3.2 节)。不幸的是，对于实际的建模工作而言，这是一个陷阱，从而阻滞了基于运动积分或作用量建模这项事业的进展[15]。对于实际星系的动力学建模，不能将 $E$ 放入分布函数的参数列表中。原因很简单：实际星系是一个多组分系统，动力学模型中至少要包括暗物质晕和恒星组分。轨道的能量不是局域量，如果 $E$ 出现在参数列表中，那么该轨道周围的相空间密度将以不受控制的方式变化。例如，最简单的轨道是恒星位于银河系中心的规则轨道，当添加暗物质晕组分后，中心引力势会变深，位于中心的恒星的能量会减少，而该轨道保持不变。但是，只有当模型完全组装好并且确定了其势能之后，我们才能知道该轨道的能量 $E_0=\Phi(0)$。因此我们无法在建模工作一开始，就指定恒星的 $f(E,\cdots)$。而当分布函数只是作用量的函数时，每个轨道周围的相空间密度则不会改变[15, 51]。

总之，相比于一般运动积分，作用量-角变量体系的优势是：作用量是一种特殊的运动积分，而且作为正则坐标，作用量有共轭坐标即角变量；作用量体系适用于使用微扰方法，也正适合描述周期性强的天体运动；作用量是绝热不变量，可用于缓慢变化的引力势，以及邻近势能中的轨道识别；此外如上段所述，使用以作用量为参数的分布函数来为星系 (多组分系统) 建模，才使得基于分布函数的建模成为可能。

### 1.3 Stäckel 势简介

Stäckel 势在椭球坐标下可分离①。三轴 Stäckel 势是一类最普遍形式的可分离势。粒子在一般 Stäckel 势中的运动，能通过代数手段得到解析的运动积分，即是完全可积的。本节介绍三轴 Stäckel 势的基本内容[1]，在这之前先回顾下椭球坐标系。

作坐标变换 $(x,y,z)\rightarrow(\lambda,\mu,\nu)$，并令

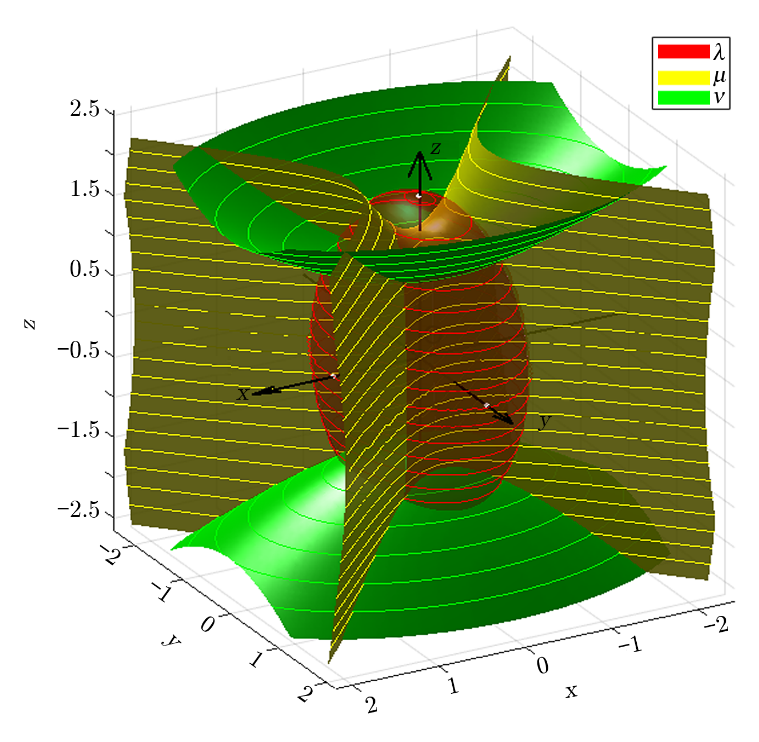
$$\frac{x^2}{\tau+\alpha}+\frac{y^2}{\tau+\beta}+\frac{z^2}{\tau+\gamma}=1\ . \tag{19}$$

其中，$\tau=\lambda,\mu,\nu$ 即构成共焦椭球坐标。令常数 $\alpha=-a^2$, $\beta=-b^2$, $\gamma=-c^2$，并且 $-\gamma\leqslant\nu\leqslant-\beta\leqslant\mu\leqslant-\alpha\leqslant\lambda$。当 $\tau$ 变化时，半轴长变化但焦距不变，形成一系列的共焦二次曲面。该三维坐标系的等坐标面如图 1 所示，如式 (19) 中令 $\tau=\lambda$，由于 $\lambda+\alpha\geqslant0$, $\lambda+\beta\geqslant0$, $\lambda+\gamma\geqslant0$，此时每个等 $\lambda$ 曲面代表一个椭球面 (图 1 中的红色面)；类似地，由于 $\mu+\alpha\leqslant0$, $\mu+\beta\geqslant0$, $\mu+\gamma\geqslant0$，等 $\mu$ 曲面代表单叶双曲面 (黄色)；等 $\nu$ 曲面代表双叶双曲面 (绿色)。笛卡尔坐标系、球坐标系等，可看作共焦椭球坐标系的极限情形。

共焦椭球坐标系也是正交坐标系，其线元 $\mathrm{d}s^2=\sum_\tau P_\tau^2\mathrm{d}\tau^2$，其度规系数为：

$$\begin{cases}P_\lambda^2=\dfrac{(\lambda-\mu)(\lambda-\nu)}{4(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)}\\[2ex] P_\mu^2=\dfrac{(\mu-\nu)(\mu-\lambda)}{4(\mu+\alpha)(\mu+\beta)(\mu+\gamma)}\\[2ex] P_\nu^2=\dfrac{(\nu-\lambda)(\nu-\mu)}{4(\nu+\alpha)(\nu+\beta)(\nu+\gamma)}\end{cases}\ . \tag{20}$$

①Stäckel 当年的工作证明，$H=\frac{1}{2}p^2+\Phi(\boldsymbol{x})$ 的哈密顿-雅克比方程分离的唯一坐标系是共焦椭球坐标系 (参考文献 [8] 第 3.5 节)。




注：椭球面 (红色) 是等 $\lambda$ 面，单叶双曲面 (黄色) 是等 $\mu$ 面，双叶双曲面 (绿色) 是等 $\nu$ 面。坐标轴数字代表长度比例。

**图 1 共焦椭球坐标系的等坐标面示意图**

此时 Stäckel 势 $\varPhi_S(\lambda,\mu,\nu)$ 可写成下面形式：

$$\begin{aligned}\varPhi_S &= -\frac{\zeta(\lambda)}{P_\lambda^2}-\frac{\eta(\mu)}{P_\mu^2}-\frac{\xi(\nu)}{P_\nu^2}\\ &= -\frac{F(\lambda)}{(\lambda-\mu)(\lambda-\nu)}-\frac{F(\mu)}{(\mu-\nu)(\mu-\lambda)}-\frac{F(\nu)}{(\nu-\mu)(\nu-\mu)}\ ,\end{aligned} \tag{21}$$

其中，$\zeta,\eta,\kappa$ 分别是相应椭球坐标 $\lambda,\mu,\nu$ 的任意函数 (见文献 [2] 中的式 (6)) [①]。令 $F_\tau(\lambda)=4(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)\zeta(\lambda)$，$F_\tau(\mu)$ 及 $F_\tau(\nu)$ 依此类推[②]，那么 Stäckel 势可以写成 3 个分量分离的形式。

将上面的 Stäckel 势代入哈密顿量

$$H=\frac{p_\lambda^2}{2P_\lambda^2}+\frac{p_\mu^2}{2P_\mu^2}+\frac{p_\nu^2}{2P_\nu^2}+\varPhi_S(\lambda,\mu,\nu)\ , \tag{22}$$

---

[①] 不同文献对引力势所采用的符号有时候不相同 ($V$、$\varPhi$ 或其他)，相应地有些文献中哈密顿量记为 $H=\frac{1}{2}p^2+V$，有些记为 $H=\frac{1}{2}p^2+\varPhi$；类似地，对其分离函数 $F_\tau(\tau)$ 前面的正负符号的规定，也可能有所出入。本文采用如式 (21) 所示的方式。

[②] 为了书写方便，本文后面统一用 $F(\tau)$ 代指分离函数 $F_\tau(\lambda)$ 等 3 个分离函数，而实际上这 3 个函数可以采取不同的具体形式。

其中，$p_\lambda = P_\lambda^2 \dot{\lambda}$，$p_\mu$ 及 $p_\nu$ 依此类推。再代入哈密顿-雅克比方程 $H = E$，以及有 $p_\tau = \partial W/\partial\tau$，其中 $W$ 为哈密顿特征函数。解哈密顿-雅克比方程，令两边乘以 $(\lambda-\mu)(\mu-\nu)(\nu-\lambda)$，可得

$$\begin{aligned}&(\nu-\mu)[2(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)\left(\frac{\partial W}{\partial\lambda}\right)^2-F(\lambda)-\lambda^2E]+\\&(\lambda-\nu)[2(\mu+\alpha)(\mu+\beta)(\mu+\gamma)\left(\frac{\partial W}{\partial\mu}\right)^2-F(\mu)-\mu^2E]+\\&(\mu-\lambda)[2(\nu+\alpha)(\nu+\beta)(\nu+\gamma)\left(\frac{\partial W}{\partial\nu}\right)^2-F(\nu)-\nu^2E]=0\ .\end{aligned} \tag{23}$$

由于式 (21) 的特殊形式以及 $W$ 的可分离性，可写为 $W=\sum_{\tau=\lambda,\mu,\nu}W_\tau(\tau)$。于是式 (23) 中括号内的部分只是各自 $\tau$ 坐标的函数，设之为 $U_\tau(\tau)$ 并替换，可得：

$$(\nu-\mu)U_\lambda(\lambda)+(\lambda-\nu)U_\mu(\mu)+(\mu-\lambda)U_\nu(\nu)=0\ . \tag{24}$$

将上式两边对每个坐标 $\boldsymbol{\tau}=\{\lambda,\mu,\nu\}$ 作两次微分，即将 $\partial^2_{\tau_1\tau_2}$ 作用于式 (24) 两边。然后我们从其中的对同坐标二次微分式中可得 $\partial^2_\tau U_\tau(\tau)=0$，即 3 个 $U_\tau(\tau)$ 是线性函数 $U_\tau(\tau)=j_\tau\tau-k_\tau$；利用不同坐标的二次微分式，可得 3 个 $U_\tau(\tau)$ 的线性系数相等，$j_\tau=j$；将 $U_\tau(\tau)$ 代入式 (24)，以及 $\{\lambda,\mu,\nu\}$ 取值的任意性 (或者再对每个坐标 $\boldsymbol{\tau}=\lambda,\mu,\nu$ 作微分)，可得 $k_\tau=k$ 相等，$k_\tau=k$，即 $j,k$ 就是可分离方程的分离常数。因此 3 个 $U_\tau(\tau)$ 只能是相同的线性函数

$$U_\tau(\tau)=j\tau-k\ . \tag{25}$$

由于 $U_\tau(\tau)$ 即式 (23) 的中括号内的部分，则可得：

$$p_\tau^2=\left(\frac{\partial W_\tau}{\partial\tau}\right)^2=\frac{\tau^2E-\tau j+k+F(\tau)}{2(\tau+\alpha)(\tau+\beta)(\tau+\gamma)}\ , \tag{26}$$

此即 Stäckel 势下动量和运动积分 $J,K$ 关系式。注意，常数 $j$，$k$ 就是除哈密顿量之外两个运动积分 $J$，$K$ 的值 (类似于总能量 $E$ 和哈密顿量 $H$ 的关系)，后面本文我们不再刻意区分；类似的还有 $i_2$，$i_3$ 之于 $I_2$，$I_3$。

更方便地，$H$，$J$，$K$ 可写为：

$$\begin{cases}H=X+Y+Z\\J=(\mu+\nu)X+(\nu+\lambda)Y+(\lambda+\mu)Z\\K=\mu\nu X+\nu\lambda Y+\lambda\mu Z\end{cases}\ , \tag{27}$$

其中，

$$\begin{cases} X = \dfrac{p_\lambda^2}{2P_\lambda^2} - \dfrac{F(\lambda)}{(\lambda-\mu)(\lambda-\nu)} \\ Y = \dfrac{p_\mu^2}{2P_\mu^2} - \dfrac{F(\mu)}{(\mu-\nu)(\mu-\lambda)} \\ Z = \dfrac{p_\nu^2}{2P_\nu^2} - \dfrac{F(\nu)}{(\nu-\lambda)(\nu-\mu)} \end{cases} . \tag{28}$$

此外，还有更常用的另一套运动积分，两者的关系是：

$$\begin{cases} I_2 = \dfrac{\alpha^2 H + \alpha J + K}{\alpha - \gamma} \\ I_3 = \dfrac{\gamma^2 H + \gamma J + K}{\gamma - \alpha} \end{cases} . \tag{29}$$

此即星系动力学中常说的“第二个、第三个运动积分” $I_2$, $I_3$。相关推导见文献 [2]。类似地，由 $p_\tau = \partial W/\partial \tau$，得到动量的表达式为

$$p_\tau^2 = \frac{(\tau+\alpha)(\tau+\gamma)E - (\tau+\gamma)I_2 - (\tau+\alpha)I_3 + F(\tau)}{2(\tau+\alpha)(\tau+\beta)(\tau+\gamma)} . \tag{30}$$

上式正是用来计算 Stäckel 势下的运动积分，也是估算一般势的运动积分时主要用到的动量表达式。由作用量的定义 (即式 (18))，可得：

$$J_\tau = \oint p_\tau \mathrm{d}\tau = \frac{2}{\pi}\int_{\tau_-}^{\tau_+} |p_\tau| \mathrm{d}\tau . \tag{31}$$

积分即得作用量，通过哈密顿特征函数则可求出角变量 $\boldsymbol{\theta}$。附录 A 介绍了如何具体数值求解作用量和角频率 (见文献 [1, 22])。

此外，Stäckel 势还有一个优势：当星系处于准稳态时，恒星分布函数的非对角二阶速度矩消失，从而仅剩的关于三个对角二阶矩的 3 个 Jeans 方程变成了闭合的；van de Ven 等人[25]用微分方程的方法给出了 Stäckel 势下 Jeans 方程的速度弥散通解。

## 2 基于 Stäckel 势近似的作用量估算方法

前面介绍了作用量-角变量的意义以及 Stäckel 势下运动积分的解析求解，然而绝大多数实际引力势下并没有解析的运动积分。接下来本章介绍几种估算一般引力势系统的作用量-角变量的估算方法，这些方法使作用量- 角变量体系得以应用于一般引力势系统。由于借用了已知的某种具体形式的 Stäckel 势作为玩具模型，因而这些方法其实蕴含了微扰法的思想。具体来说，这些方法的过程是：给定某种已知的具体形式的 Stäckel 势作为玩具模型势，并在一定程度上用所画的“真实”轨道上的采样点数据来拟合相应的待定模型引力势 (或者生成函数里的参数)；然后用这个计算的势作为“真实 Stäckel 势”来估算运动积分，从而计算作用量及角频率等所需要的量。

Sanders 和 Binney[22, 24]还提供了相应的开源程序 TACT ①。该程序用 C++ 所写，能够用若干种方法计算出作用量、角频率和角变量，并能给出误差；另外，还提供了相应的 Python 画图程序。关于算法的更多细节，请参见文献 [20, 24]，以及 TACT 程序的源码。

## 2.1 Stäckel 作弊法

关于估算一般引力势系统的作用量-角变量，Binney[13]提出了高效的 Stäckel Fudge 方法，可看作是一种作弊算法 (以后称为 Stäckel 作弊法)。其思想是：假设引力势可以局部地写成 Stäckel 势的形式，那么以一些技巧局部地 (即在恒星轨道的某一段之内) 计算目标势的伪分离函数 (参见式 (21) 的 $F(\tau)$)，并且假定其他坐标对分离函数误差的交叉影响较小，然后以此计算一组新的运动积分，得到动量，从而通过积分得到作用量。下面先简介 Binney 首先提出的轴对称情形的 Stäckel 作弊法，然后介绍三轴 Stäckel 作弊法[22]，后者是对前者的扩展。

### 2.1.1 轴对称势的 Stäckel 作弊法

一般来说，轴对称势下有 3 个孤立运动积分，即总能量 $E$，角动量垂向分量 $L_z$，以及 $I_3$。轴对称椭球坐标系相对于三轴椭球坐标系而言，一个半轴长参数与另外一个相等，如果 $\alpha=\beta$，则为长等轴 (prolate) 椭球坐标系 $(\lambda,\phi,\nu)$，其中 $\phi$ 代表某个子午面的方位角；如果 $\gamma=\beta$，则为扁等轴 (oblate) 椭球坐标系 $(\lambda,\mu,\chi)$，其中 $\chi$ 代表某个子午面的方位角。因此在描述粒子在轴对称势中的运动时，可以消去一个自由度。

基于等轴椭球坐标系作出新的 $(u,v,\phi)$ 坐标 (参考文献 [8] 第 3.5.3 节)，其中 $R=\Delta\sinh u\sin v$, $z=\Delta\cosh u\cos v$。该坐标系下轴对称 Stäckel 势为：

$$\Phi_S(u,v)=\frac{F_u(u)-F_v(v)}{\sinh^2 u+\sin^2 v}\ . \tag{32}$$

由类似于式 (30) 的关系，需要显式的 Stäckel 分离函数 $F_\tau(\tau)$，才能得到用于计算作用量的动量 $p_\tau^2$。

若目标引力势 $\Phi(u,v)$ 写成Stäckel势的形式，则 $F_u(u)-F_v(v)\approx(\sinh^2 u+\sin^2 v)\Phi_S(u,v)$；并且若目标引力势接近 Stäckel 势，则 $\delta F_u(u)$ 对 $v$ 的依赖和 $\delta F_v(v)$ 对 $u$ 的依赖是很小的。于是：

$$\delta F_u(u)\equiv(\sinh^2 u+\sin^2 v)\Phi(u,v)-(\sinh^2 u_0+\sin^2 v)\Phi(u_0,v)\approx F_u(u)-F_u(u_0)\ , \tag{33}$$

同样，$F_v(v)$ 可以重写为 $F_v(\pi/2)+\delta F_v(v)$。

这里 $u_0$ 是一个不影响作用量计算的参考值，其选取是固定 $v$ 而使 $\delta F_u(u)$ 最小值时的 $u$ (见文献 [13])。$F_u(u_0)$ 和 $F_v(\pi/2)$ 的计算为：

① https://github.com/jls713/tact

$$\begin{cases} I_3+F_u(u_0)\approx I_3+F_u(u)-\delta F_u(u)=E\sinh^2 u-\dfrac{p_u^2}{2\Delta^2}-\dfrac{L_z^2}{2\Delta^2\sinh^2 u}-\delta F_u(u) \\ I_3+F_v(\pi/2)\approx I_3+F_v(v)-\delta F_v(v)=-E\sin^2 v+\dfrac{p_v^2}{2\Delta^2}+\dfrac{L_z^2}{2\Delta^2\sinh^2 v}-\delta F_v(v) \end{cases} . \tag{34}$$

2.1.2 三轴势的 Stäckel 作弊法

本小节主要介绍 Sanders 和 Binney[22] 提出的算法，其是对 Binney[13] 提出的 Stäckel 作弊法的三轴推广。对于一般轴对称势系统，除了 Stäckel 作弊法，Sanders[19] 还提供了对一般轴对称势的 Stäckel 拟合的方法 (下一小节介绍)。但是三轴 Stäckel 势比轴对称的多一维参数，对其拟合所需的具体拟合模型会复杂许多，于是 Sanders 和 Binney 推广出了三轴势的 Stäckel 作弊法。该方法借助类似于运动积分 $J$, $K$ 的另一套运动积分 (作弊法得到的新的不独立的运动积分)，以及局部近似的 Stäckel 势，用动量和运动积分的关系式得到每一点的动量 $p_\tau$，再积分得到作用量 $\boldsymbol{J}$；类似地，可求出角变量 $\boldsymbol{\theta}$。本方法以当前真实的相空间点做初始条件 $(x_0, v_0)$，来计算一系列动量，这样就能不需要 Stäckel 近似势的分离势 $F(\tau)$ 的显式表达式；然后积分得到作用量；最小化作用量的标准差 (与每个初始 $(x_0, v_0)$ 有关)，调整焦距，从而再得到合适的作用量；对于轴对称势系统，还可做迭代以减小误差。下面介绍其具体步骤。

(1) 三轴势 Stäckel 作弊法与运动积分的计算

对于一个一般引力势，定义三轴情形的辅助函数：

$$\begin{cases} \chi_\lambda(\lambda,\mu,\nu)=(\lambda-\mu)(\nu-\lambda)\varPhi(\lambda,\mu,\nu) \\ \chi_\mu(\lambda,\mu,\nu)=(\mu-\nu)(\lambda-\mu)\varPhi(\lambda,\mu,\nu) \\ \chi_\nu(\lambda,\mu,\nu)=(\nu-\lambda)(\mu-\nu)\varPhi(\lambda,\mu,\nu) \end{cases} , \tag{35}$$

如果引力势就是 Stäckel 势，那么式 (35) 可以写为：

$$\chi_\lambda(\lambda,\mu,\nu)=F(\lambda)-\lambda\frac{F(\mu)-F(\nu)}{\mu-\nu}+\frac{\nu F(\mu)-\mu F(\nu)}{\mu-\nu} . \tag{36}$$

现在，假设该引力势是近似于 Stäckel 形式的，从而由上式可以写成：

$$F(\tau)=\chi_\tau(\lambda,\mu,\nu)+\tau C_\tau+D_\tau . \tag{37}$$

当固定另外两个椭球坐标时，某个 $\tau$ 对应的 $C_\tau, D_\tau$ 可看作常数，从而使我们可以分开来独立计算所需的运动积分。

将式 (35) 代入到式 (26)，得到动量的另一个计算式

$$2(\tau+\alpha)(\tau+\beta)(\tau+\gamma)p_\tau^2=\tau^2E-\tau J_\tau+K_\tau+\chi_\tau(\lambda,\mu,\nu) . \tag{38}$$

与式 (37) 相对比，就有了这些常数之间的关系式 $J_\tau=j-C_\tau$，$K_\tau=k+D_\tau$ ($j$, $k$ 正是前面式 (26) 的运动积分)。这样，给定一个初始相空间点 ($\boldsymbol{x}_0$, $\boldsymbol{v}_0$) 和一套合适的椭球坐标系 (坐

标变换参数即 $(\alpha,\beta,\gamma)$)，以及引力势各点的值，便可计算某椭球坐标的相空间点时的运动积分[22]。

下面介绍 $J_\tau$, $K_\tau$ 的计算。由式 (38) 得：

$$K_\tau = 2(\tau_0+\alpha)(\tau_0+\beta)(\tau_0+\gamma)p_{\tau_0}^2 - \tau_0^2 E + \tau_0 J_\tau - \chi_\tau(\lambda_0,\mu_0,\nu_0) \ . \tag{39}$$

$K_\tau$ 依赖于 $J_\tau$，而其他量可以计算。当固定另外两组椭球坐标和动量时，由于 Stäckel 势的分离性，可以通过仅改变某个 $\tau$ 和 $p_\tau$ 的值，便可使粒子仍保持在既定轨道上，从而认为在 Stäckel 势下保持椭球相空间坐标不变，哈密顿量对某个椭球坐标的偏导为 0 (见文献 [22] 中的第 3 章)，即：

$$0 = (\frac{\partial H}{\partial\lambda})_{\mu,\nu} = \frac{1}{2}\frac{\partial}{\partial\lambda}\left[\frac{p_\lambda^2}{P_\lambda^2}\right] + \frac{1}{2}\frac{p_\mu^2}{(\mu-\lambda)P_\mu^2} + \frac{1}{2}\frac{p_\nu^2}{(\nu-\lambda)P_\nu^2} + \frac{\partial\Phi}{\partial\lambda} \ . \tag{40}$$

分别将 $\lambda$ 和 $\lambda_0$ 代入式 (39)，并作差，除以 $P_\lambda$，可推得

$$\frac{1}{2}\frac{p_\lambda^2}{P_\lambda^2} = \frac{[2(\lambda_0+\alpha)(\lambda_0+\beta)(\lambda_0+\gamma)p_{\lambda_0}^2 + \chi_\lambda(\lambda_0,\mu_0,\nu_0)] + (\lambda^2-\lambda_0^2)E - (\lambda-\lambda_0)J_\lambda}{(\lambda-\mu)(\lambda-\nu)} - \Phi(\lambda,\mu_0,\nu_0) \ . \tag{41}$$

用式 (41) 来计算式 (40) 第一项，可见同时可以消去 $\partial\Phi/\partial\lambda$，因此得到 $J_\lambda$ 以及类似的 $J_\mu$, $J_\nu$ 的表达式：

$$\begin{cases} J_\lambda = (\mu_0+\nu_0)E + \dfrac{1}{2}\dfrac{p_{\mu_0}^2(\lambda_0-\mu_0)}{P_{\mu_0}^2} + \dfrac{1}{2}\dfrac{p_{\nu_0}^2(\lambda_0-\nu_0)}{P_{\nu_0}^2} \\ J_\mu = (\lambda_0+\nu_0)E + \dfrac{1}{2}\dfrac{p_{\lambda_0}^2(\mu_0-\lambda_0)}{P_{\lambda_0}^2} + \dfrac{1}{2}\dfrac{p_{\nu_0}^2(\mu_0-\nu_0)}{P_{\nu_0}^2} \\ J_\nu = (\lambda_0+\mu_0)E + \dfrac{1}{2}\dfrac{p_{\lambda_0}^2(\nu_0-\lambda_0)}{P_{\lambda_0}^2} + \dfrac{1}{2}\dfrac{p_{\mu_0}^2(\nu_0-\mu_0)}{P_{\mu_0}^2} \end{cases} . \tag{42}$$

于是得到 6 个估算的运动积分 $J_\tau$, $K_\tau$ 的表达式，其中只有 3 个是独立的[22]。

(2) 合适焦距的选取

有了这些新的运动积分，可用式 (38) 求得相应 $\tau$ 点的动量，之后积分得到作用量。轴对称情形也类似。然而，坐标变换的参数的焦距 ($\Delta_1=\sqrt{\beta-\alpha}$, $\Delta_2=\sqrt{\gamma-\beta}$，任取 $\gamma=-1\,\mathrm{kpc}^2$)，其大小非常影响该方法所估算的作用量的方差 (认为是误差)。Sanders 和 Binney 给出了数值调整合适焦距的方法。例如对于短轴闭环轨道 (short axis closed loop orbits)，轨道形成在 $z=0$ 面，焦距为 $y=\pm\Delta_1$ 的椭圆。在这种轨道，只有作用量 $J_\nu$ 不为 0，利用这一点，给每个能量值找出一个合适的焦距。

其步骤是，先给定一个一般势，最大和最小能量值以及期间的能量网格点；在每个网格点 $E_i$，画一个从坐标 $(0,y_k,0)$ 以速度为 $\sqrt{2(E_i-\Phi(0,y_k,0))}$ 垂直出发的半周期的轨道；记录所画轨道的相空间点并计算作用量 $J_\nu$，求出使作用量的方差最小的焦距 $\Delta_1$。$\Delta_2$ 的计算也可用上述方法求出。

(3) 环面迭代增加精度

Stäckel 作弊法的过程复杂，且误差较大；偏差范围可达百分之几至几十，特别是盒轨道的偏差更大 (见文献 [22] 中的图 6−8)。但是若把它与环面映射 (见 2.3.1 节) 方法相结合，通过环面迭代，可使误差收敛。这种迭代算法将在 2.3.2 节具体介绍。

## 2.2 Stäckel 拟合法

### 2.2.1 一般轴对称势的 Stäckel 拟合法

这里介绍 Sanders[19] 提出的长等轴椭球坐标系下的一般轴对称势的 Stäckel 拟合法。该方法是对一般引力势进行 Stäckel 形式的拟合，即采样轨道来拟合出一个合适的 Stäckel 引力势的分离函数 $F(\tau)$，来得到运动积分。Sanders 和 Binney[24] 把 Stäckel 作弊法和这种 Stäckel 拟合法归类为“非收敛方法”，因为这两种方法依赖于做一定的近似，而近似所导致的误差不能通过增加计算量来不断减小。

在某个子午面上，运动的位形空间从柱坐标 $(R,\phi,z)$ 变换到扁球坐标 $(\lambda,\phi,\nu)$ 依赖于

$$\frac{R^2}{\tau+\alpha}+\frac{z^2}{\tau+\gamma}=1 \ , \tag{43}$$

其中，$\lambda$, $\nu$ 是上式作为关于 $\tau$ 的方程根，$a^2=\alpha$, $b^2=\beta$, $c^2=\gamma$，类似于三轴情形下的式 (19)。此时 Stäckel 势的形式简化为：

$$\varPhi_S=-\frac{F(\lambda)-F(\nu)}{\lambda-\nu} \ . \tag{44}$$

轴对称时第三运动积分的解析表达式为：

$$I_3=(\lambda+\gamma)\left[E-\frac{L_z^2}{2(\lambda+\alpha)}+\frac{F(\lambda)}{\lambda+\gamma}-\frac{\dot{\lambda}^2(\lambda-\nu)^2}{8(\lambda+\alpha)(\lambda+\gamma)^2}\right] \ . \tag{45}$$

动量和运动积分的关系表示为：

$$p_\tau^2=\frac{1}{2(\tau+\alpha)}\left[E-\frac{L_z^2}{2(\lambda+\alpha)}+\frac{F(\tau)}{\tau+\gamma}-\frac{I_3}{\tau+\gamma}\right] \ . \tag{46}$$

下面具体介绍对一般轴对称势的 Stäckel 拟合和环面坐标建立的方法以及每步所需要的手续。

(1) 给定初始的 $(x,v)$，以及真实的势 (轨道积分中所感受到的势)，并在真实的势中计算实际的运动积分 $E, L_z$。

(2) 焦距的选取。作上述变换到椭球坐标。然而，变换的参数，即焦距 $\varDelta=a^2-c^2$ (因为 $a=b$，只要一个 $a^2-c^2$ 便可确定椭球坐标系，可取 $c^2=1$)，非常影响作用量的标准差，因此通过一定的手续来估算一个更好的焦距 $\varDelta$。Sanders[19] 采用的估计方法是，由于 Stäckel 势的分离性，可得

$$\frac{\partial^2}{\partial\lambda\partial\nu}((\lambda-\nu)\varPhi_S)=0 \ . \tag{47}$$

由上式以及两种坐标之间的变换关系，可推得[19]：

$$a^2 - c^2 = R^2 - z^2 - \left[3z\frac{\partial \Phi_S}{\partial R} - 3R\frac{\partial \Phi_S}{\partial z} + Rz\left(\frac{\partial^2 \Phi_S}{\partial R^2} - \frac{\partial^2 \Phi_S}{\partial z^2}\right)\right] / \frac{\partial^2 \Phi_S}{\partial R \partial z} \ . \tag{48}$$

既然假设所估算的局部势 $\Phi$ 就是该 Stäckel 的势 $\Phi_S$，于是就用该式来估计 $a^2$ 的值，其中的偏导用已知的真实 $\Phi$ 来数值计算梯度而得；每个 $a^2$ 值，需要在轨道中每个点的相应值来计算，最后做平均，这些点在下一步的轨道采样中获得。

(3) 轨道采样以及运动区域的估算。进行轨道积分，即在 $(x, v)$ 坐标下通过每一点的势和时间积分来算下一步的坐标 $(x',\ v')$，即相当于数值法画轨道。Sanders 和 Binney 的 TACT 程序中用守恒格式的数值方法画出轨道，并且每个轨道周期离散为 300 个点。

在画轨道时：1) 在前几周期的一些点中，隔 5 个点用式 (48) 估算 $a^2$ ($c^2$ 已取 1) 并平均，最后得到那个最合适的扁球坐标；2) 通过前后 $\partial\tau/\partial t$ 异号等条件 (其中每次需要 $x$ 到 $\tau$ 的坐标变换) 来判断是否达到轨道的 $\tau$ 的端点，找到后就记录，得到 $\lambda_{+-}, \nu_+, \nu_- = c^2$)，从而得到轨道在椭球坐标的所在区域。

(4) 样条插值拟合 Stäckel 势。关于引力势拟合，这里假设局部是 Stäckel 的，于是先在椭球坐标下若干点计算出相应的“假设是 Stäckel 形式的所求势”的值；之后对这些点的势的值做三次样条插值得到任意点的值。Stäckel 辅助函数 $\chi$ 在轴称下简化为 $\chi = -(\lambda, \nu)\Phi(\lambda, \nu)$。

假如所拟合的势 $\Phi$ 是 Stäckel 的时，会有 $-(\lambda, \nu)\Phi(\lambda, \nu) = F(\lambda) - F(\nu)$，正是据此得到若干插值点局部的假设是 Stäckel 形式的所求引力势的值。关于 $F(\tau)$ 的选取，其要求也正是使其与 Stäckel 势的偏差

$$E[F] = \int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda \int_{\nu_-}^{\nu_+} \mathrm{d}\nu \Lambda(\lambda) N(\nu) (\chi(\lambda, \nu) - F(\lambda) + F(\nu))^2 \tag{49}$$

最小。归一化的权重函数形式为：

$$\Lambda(\lambda) = 4\lambda^{-5}(\lambda_-^{-4} - \lambda_+^{-4})^{-1} \ , \quad N(\nu) = (\nu_+ - \nu_-)^{-1} \ . \tag{50}$$

这个权重函数是提前选的，相关内容可见文献 [1, 4]。于是拟合中所选用的 Stäckel 势的具体形式为：

$$F(\lambda) = \bar{\chi}(\lambda) - \frac{1}{2}\bar{\bar{\chi}} \ , \quad F(\nu) = -\bar{\chi}(\nu) + \frac{1}{2}\bar{\bar{\chi}} \ , \tag{51}$$

其中，

$$\begin{aligned} \bar{\chi}(\lambda) &= \int_{\nu_-}^{\nu_+} \mathrm{d}\nu \chi(\lambda, \nu) N(\nu) \ , \quad \bar{\chi}(\nu) = \int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda \chi(\lambda, \nu) \Lambda(\lambda) \ , \\ \bar{\bar{\chi}} &= \int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda \int_{\nu_-}^{\nu_+} \mathrm{d}\nu \chi(\lambda, \nu) \Lambda(\lambda) N(\nu) \ . \end{aligned} \tag{52}$$

这是因为在这种情况下偏差是最小的。证明方法是，式 (49) 取最小时，对于 $\lambda$, $\nu$ 分别成立，

即将其分别对 $\lambda,\nu$ 求导应为 0，得

$$\begin{cases} \displaystyle\int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda \Lambda(\lambda)[\chi(\lambda,\nu) - F(\lambda) + F(\nu)] = 0 \ , \\ \displaystyle\int_{\nu_-}^{\nu_+} \mathrm{d}\nu N(\nu)[\chi(\lambda,\nu) - F(\lambda) + F(\nu)] = 0 \ . \end{cases} \tag{53}$$

再做一些代换即可得到式 (50)。

接下来，有了式 (50) 的形式后，便可结合式 (51) 和 (52) 直接积分得到某个点 $(\lambda,\ \nu)$ 的 $F(\tau)$，如此选取若干个坐标点 (TACT 程序选取了 40 个点)，用来三次样条插值得到所拟合的势 (样条插值使拟合的势会更平滑)。

(5) 计算运动积分。借用这个拟合势，以及之前的 $E,\ L_z$，用第三运动积分的式 (45) 在 3 个端点 $\tau_{+-}$ 计算 $I_3$ 并平均，从而得到 3 个相应运动积分。其中计算时需用真实性条件 $p_\tau^2 > 0$ 进行判断，如果不符合，就不用端点而用初始相空间点。

(6) 计算作用量-角变量。计算的 $F(\tau)$ 除了上一步用来算 $I_3$，还用来在式 (46) 中以此得到当前点的 $p_\tau$，从而计算作用量 $J$ 和角变量 $\theta$ (数值计算步骤具体见附录 A)。

#### 2.2.2 Stäckel 拟合的其他相关方法

早在 1985 年，de Zeeuw 和 Lynden-Bell[2] 就已提出了把 Stäckel 势作为模型，来局部拟合和全局拟合星系引力势的思想和方法。前者在具体星系引力势的平衡点附近展开引力势，然后通过对应 Stäckel 的椭球坐标形势来确定参数；后者通过以平均的方法来近似 Stäckel 势的 3 个分离函数，从而计算运动积分。下面简单介绍。

(1) 局部拟合

在平衡位置附近将 $V = V(x^2, y^2, z^2)$ 形式的引力势展开为：

$$V = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} V_{klm} x^k y^l z^m \quad , \tag{54}$$

其中，指标 $k,\ l,\ m$ 只取偶数值。另外，Stäckel 势可写为式 (21) 的分离形式，以 $(\lambda,\mu,\nu) = (-\alpha,-\beta,-\gamma)$ 作为展开源点，展开分离的函数选取在展开源点附近的引力势为 0，即 $\zeta_{-1} = \eta_{-1} = \kappa_{-1} = 0$。于是可得：

$$\zeta(\lambda) = \sum_{n=0}^{\infty} \zeta_n (\lambda+\alpha)^n \ , \quad \eta(\mu) = \sum_{n=0}^{\infty} \eta_n (\mu+\beta)^n \ , \quad \kappa(\nu) = \sum_{n=0}^{\infty} \kappa_n (\nu+\gamma)^n \ . \tag{55}$$

推导中用到了 $(\lambda+\alpha)$ 的展开式，比如：

$$\lambda + \alpha = x^2 \left\{ 1 + \frac{y^2}{\beta-\alpha} + \frac{z^2}{\gamma-\alpha} - \frac{y^2(x^2-y^2)}{(\beta-\alpha)^2} + \frac{2y^2z^2}{(\beta-\alpha)(\gamma-\alpha)} - \frac{z^2(x^2-z^2)}{(\gamma-\alpha)^2} \cdots \right\} \ . \tag{56}$$

$P_\lambda^2$ 也同样地展开。对比式 (21) 的形式，式 (54) 的展开系数应为

$$\begin{aligned} &V_{000}=0\ ,\quad V_{200}=-4\zeta_0\ ,\quad V_{400}=-4\zeta_1\ ,\quad V_{220}=-\frac{8}{\beta-\alpha}(\zeta_0-\eta_0)\ ,\\ &V_{600}=-4\zeta_2\ ,\quad V_{420}=-\frac{12}{(\beta-\alpha)^2}[\zeta_1(\beta-\alpha)-\zeta_0+\eta_0]\ ,\\ &V_{222}=\frac{24[\zeta_0(\gamma-\beta)+\eta_0(\alpha-\gamma)+\kappa_0(\beta-\alpha)]}{(\gamma-\beta)(\alpha-\gamma)(\beta-\alpha)}\ . \end{aligned} \tag{57}$$

文献 [2] 还讨论了这些系数的约束条件、存在性和合适形式的选取。

(2) 全局拟合

文献 [2] 的第 4 章对椭球势做了讨论，表明与椭圆星系有关的引力势不仅可以局部地，而且还可以在全局地近似为 Stc̆kel 形式，并且给出了椭球势 (具体为镜像对称的引力势 $V=V(x^2,y^2,z^2)$) 的全局拟合方法。该方法要求所拟合的势在整体上离 Stäckel 势最近，并且当 $V$ 恰好是 Stäckel 势时应得到精确的结果。下面列出其方法的主要步骤。

对于任意函数 $Q=Q(\lambda,\mu,\nu)$ 定义下列平均

$$\begin{cases} \langle Q,\lambda\rangle=\dfrac{1}{\varLambda MN}\displaystyle\int_{-\beta}^{-\alpha}\int_{-\gamma}^{-\beta}Q(\lambda,\mu,\nu)M(\mu)N(\nu)\mathrm{d}\mu\mathrm{d}\nu\\ \langle Q,\mu\rangle=\dfrac{1}{\varLambda N}\displaystyle\int_{-\alpha}^{-\infty}\int_{-\gamma}^{-\beta}Q(\lambda,\mu,\nu)\varLambda(\lambda)N(\nu)\mathrm{d}\lambda\mathrm{d}\nu\\ \langle Q,\nu\rangle=\dfrac{1}{M\varLambda}\displaystyle\int_{-\alpha}^{-\infty}\int_{-\beta}^{-\alpha}Q(\lambda,\mu,\nu)\varLambda(\lambda)M(\mu)\mathrm{d}\lambda\mathrm{d}\mu\\ \bar{Q}=\dfrac{1}{\varLambda MN}\displaystyle\int_{-\alpha}^{-\infty}\int_{-\beta}^{-\alpha}\int_{-\gamma}^{-\beta}Q(\lambda,\mu,\nu)\varLambda(\lambda)M(\mu)N(\nu)\mathrm{d}\lambda\mathrm{d}\mu\mathrm{d}\nu \end{cases}\ , \tag{58}$$

其中，$\varLambda(\lambda)$, $M(\mu)$, $N(\nu)$ 是根据积分的约束选择的权重函数，相应的归一化常数为：

$$\varLambda=\int_{-\alpha}^{\infty}\varLambda(\lambda)\mathrm{d}\lambda\ ,\quad M=\int_{-\beta}^{-\alpha}M(\mu)\mathrm{d}\mu\ ,\quad N=\int_{-\gamma}^{-\beta}N(\nu)\mathrm{d}\nu\ . \tag{59}$$

定义辅助函数

$$\chi(\lambda,\mu,\nu)=-(\lambda-\mu)(\mu-\nu)(\nu-\lambda)V(\lambda,\mu,\nu)\ . \tag{60}$$

那么，分离函数可计算为：

$$\begin{cases} F(\lambda)=-\dfrac{\langle\chi-2/3\bar{\chi},\lambda\rangle}{(\bar{\mu}-\nu)}-\dfrac{\bar{\chi}\cdot(\lambda-\bar{\lambda})(2\bar{\lambda}-\bar{\mu}-\bar{\nu})}{3L}\\ F(\mu)=-\dfrac{\langle\chi-2/3\bar{\chi},\mu\rangle}{(\bar{\nu}-\bar{\lambda})}-\dfrac{\bar{\chi}\cdot(\mu-\bar{\mu})(2\bar{\mu}-\bar{\nu}-\bar{\lambda})}{3L}\\ F(\nu)=-\dfrac{\langle\chi-2/3\bar{\chi},\nu\rangle}{(\bar{\lambda}-\bar{\mu})}-\dfrac{\bar{\chi}\cdot(\nu-\bar{\nu})(2\bar{\nu}-\bar{\lambda}-\bar{\mu})}{3L} \end{cases}\ , \tag{61}$$

其中，$L=(\bar{\lambda}-\bar{\mu})(\bar{\mu}-\nu)(\bar{\nu}-\bar{\lambda})$。此即文献 [2] 中的式 (43)，更多的约束细节可见该文。

尽管相比于近年改进的新方法 (前述的 Stäckel 作弊法和 Stäckel 拟合法)，这些早期方法的适用性较小，但可以说，其基本思想是相似的或者被沿用了。例如，上小节所述的 Stäckel 拟合法，就与这里的拟合思路一致：把 Stäckel 势作为模型，来拟合具体星系的引力势。而 Stäckel 作弊法与 Stäckel 拟合法的区别仅在于：Stäckel 拟合法需要计算出显式的分离函数 $F(\tau)$，即局部地基于某些采样点把给定的引力势拟合为 Stäckel 形式，并计算分离函数，然后对这些采样点上的分离函数进行插值得到最佳拟合的 Stäckel 势，之后就依照 Stäckel 势理论中的各种解析公式来计算运动积分及作用量；而 Stäckel 作弊法是把给定的引力势局部地当作 Stäckel 势，即假设给定引力势的局部可分离性，因此可以采用一定的技巧来计算局部的某种运动积分，之后就是同样地依照 Stäckel 势理论中的各种解析公式来计算作用量。

### 2.3 基于环面映射的方法

本节介绍的内容包括：(1) McGill 和 Binney[11] 提出、Binney 及其合作者后来完善的，通过正则变换构造环面 (作用量-角变量) 来计算作用量的方法，即基于玩具模型的环面映射 (torus mapping, TM) 思想及其 TM 程序[12]；(2) 基于此思想的两个扩展：迭代环面构造法 (iterative torus construction, ItTC)[22, 24] 和轨道积分拟合生成函数法 (orbit integration to generation function, O2GF)[21, 24]。这类方法的特点是精度最高，计算花费大，属于收敛方法，即可以通过充分大的计算量，把作用量的误差一直不断地减小下去。

将环面 (n-Torus) 视为几何物体，而将解析的哈密顿量的环面变形为所求的目标哈密顿量的环面，如同微扰理论，通过寻求从玩具环面坐标到目标环面坐标的正则变换来实现[11]。即用已知的解析引力模型作为玩具模型，来逼近一般的目标引力系统，而这是在环面坐标的意义上进行的，并且通过正则变换来实现，此即环面映射的关键思想。一般所用的玩具模型为等时势或谐振子势，如式 (70) 和 (71)，这些都是 Stäckel 势，可解析求解。尽管环面映射法不同于前述的通过近似 Stäckel 的分离函数 $F_\tau(\tau)$ 的 Stäckel 作弊法，但其需要从某种作为玩具模型的 Stäckel 势出发，本质上相当于某种简化版的微扰法，因而仍可看作是 Stäckel 势的应用拓展。

本节中，若无特殊说明，用 $(\boldsymbol{J}^{\rm toy}, \boldsymbol{\theta}^{\rm toy})$ 代表玩具环面 (即玩具势下粒子运动的环面坐标)，$(\boldsymbol{J}, \boldsymbol{\theta})$ 代表目标环面 (即所求的一般引力势下粒子运动的环面坐标)。从玩具环面映射到目标环面，将此看作一个正则变换，选取第二类生成函数 $S = S(\boldsymbol{J}, \boldsymbol{\theta}^{\rm toy})$，将这个生成函数对角变量傅里叶展开，其一般形式为：

$$S(\boldsymbol{\theta}^{\rm toy}, \boldsymbol{J}) = \boldsymbol{\theta}^{\rm toy} \cdot \boldsymbol{A}(\boldsymbol{J}) - {\rm i}\sum_{\boldsymbol{n}\neq\boldsymbol{0}} S_{\boldsymbol{n}}(\boldsymbol{J}) \exp({\rm i}\boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy}) \ , \tag{62}$$

其中，这里的虚单位 i 是为后面求导表达式的简便而加的。当沿着玩具环面在 $\boldsymbol{\theta}$ 的任一分量增加一个周期 $2\pi$，相应地，目标圆环上的像点也应完成一个周期回路，最终这要求 $\boldsymbol{A} = \boldsymbol{J}$。于是

$$S(\boldsymbol{\theta}^{\rm toy}, \boldsymbol{J}) = \boldsymbol{\theta}^{\rm toy} \cdot \boldsymbol{J} - {\rm i}\sum_{\boldsymbol{n}\neq\boldsymbol{0}} S_{\boldsymbol{n}}(\boldsymbol{J}) \exp({\rm i}\boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy}) \ . \tag{63}$$

通过作用量-角变量的实数性以及哈密顿量的时间反演不变性，可以对生成函数的系数进行

约束[11]：

$$S_{\boldsymbol{n}} = -S_{\boldsymbol{n}} \ . \tag{64}$$

于是有[21]：

$$S(\boldsymbol{\theta}^{\rm toy}, \boldsymbol{J}) = \boldsymbol{\theta}^{\rm toy} \cdot \boldsymbol{J} + 2 \sum_{\boldsymbol{n} \in \boldsymbol{N}} S_{\boldsymbol{n}}(\boldsymbol{J}) \sin \boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy}, \tag{65}$$

其中，向量 $\boldsymbol{N} = \{(i,\ j,\ k)\}$ 代表一系列精度的坐标 $(i,\ j,\ k)$ 集，$(i,\ j,\ k)$ 满足 $(k > 0)$ 或 $(k = 0,\ j > 0)$ 或 $(k = 0,\ j = 0,\ i > 0)$[12, 21]。对这个生成函数 $S(\boldsymbol{\theta}^{\rm toy},\ \boldsymbol{J})$ 求微分，得

$$\boldsymbol{J}^{\rm toy} = \frac{\partial S}{\partial \boldsymbol{\theta}^{\rm toy}} = \boldsymbol{J} + 2 \sum_{\boldsymbol{n} \in \boldsymbol{N}} \boldsymbol{n} S_{\boldsymbol{n}}(\boldsymbol{J}) \sin \boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy} \ , \tag{66}$$

$$\boldsymbol{\theta} = \frac{\partial S}{\partial \boldsymbol{J}} = \boldsymbol{\theta}^{\rm toy} + 2 \sum_{\boldsymbol{n} \in \boldsymbol{N}} \frac{\partial S_{\boldsymbol{n}}(\boldsymbol{J})}{\partial \boldsymbol{J}} \sin \boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy} \ . \tag{67}$$

这样就通过生成函数联系起了玩具环面和目标环面。

#### 2.3.1　环面映射程序

Binney 和 McMillan[12]提供了 TM 程序 (Torus Mapper)①，该程序实现三个方面的功能，具体功能与输入输出如下：用户写出目标引力势及其一阶微分，以及给定作用量 $(J_r, J_\phi, J_z)$，1) 对于特定 $\boldsymbol{\theta}$，可计算 $(\boldsymbol{x},\ \boldsymbol{v})$；2) 计算雅克比 $\partial(\boldsymbol{x})/\partial(\boldsymbol{\theta})$；3) 给定 $(\boldsymbol{x}, \boldsymbol{J})$，计算 $\boldsymbol{\theta}$。与传统的恒星轨道相比，环面的优点是：一个环面是由 100 个数字指定的，即 $S_{\boldsymbol{n}}$，$\partial S_{\boldsymbol{n}}/\partial \boldsymbol{J}$ 和一些其他参数，而不是由沿时间序列的数千个相空间坐标指定，因此可以在给定分辨率下显著缩小轨道库的大小[12]。

环面映射方法的流程见式 (68)[12]，通过三步正则变换完成，变换形式为：

$$(\boldsymbol{\theta}, \boldsymbol{J}) \xrightarrow{S_{\boldsymbol{n}}} (\boldsymbol{\theta}^{\rm toy}, \boldsymbol{J}^{\rm toy}) \xrightarrow{\text{哈雅方程}} (\boldsymbol{x}^{\rm toy}, \boldsymbol{v}^{\rm toy}) \xrightarrow{\text{点变换}} (\boldsymbol{x}, \boldsymbol{v}) \ . \tag{68}$$

已知的真实势和轨道，给定目标作用量 $\boldsymbol{J}$ 和系数 $S_{\boldsymbol{n}}$ 的试探值，由玩具坐标 $\boldsymbol{\theta}^{\rm toy}$的网格点，根据式 (66) 可得到相应的玩具作用量 $\boldsymbol{J}^{\rm toy}$。在玩具势下，可直接通过解析的方法从环面坐标得到相空间坐标 $(\boldsymbol{x}^{\rm toy},\ \boldsymbol{v}^{\rm toy})$。有了这些相空间坐标可计算相应采样点的能量，之后用 Levenberg-Marquardt 算法迭代以最小化哈密顿量偏差 $\delta H$，得到最终的 $S_{\boldsymbol{n}}$。类似地，可得到 $\partial S_{\boldsymbol{n}}/\partial \boldsymbol{J}$，再通过式 (67) 可得到目标角变量 $\boldsymbol{\theta}$[24]。

这样，通过真实的环面坐标 (目标环面坐标作为已知)，用合适的系数 $S_{\boldsymbol{n}}(\boldsymbol{J})$ 和 $\partial S_{\boldsymbol{n}}/\partial \boldsymbol{J}$，转化到玩具环面坐标；再通过玩具势下的解析性就得到相空间坐标 $(\boldsymbol{x}^{\rm toy},\ \boldsymbol{v}^{\rm toy})$，并由点变换得到 $(\boldsymbol{x},\ \boldsymbol{v})$。点变换也是一类正则变换，在程序中仅当 $J_z/J_r > 0.05$ 且尝试在不进行点变换的情况下拟合环面失败时，再使用点变换[12]。

文献 [12] 中的附录图 A1 给出了环面映射方法的原理示意，对于一个好的玩具模型，更高阶的生成函数傅里叶系数能越来越好地拟合轨道。

①TM程序的源码可下载于：https://github.com/PaulMcMillan-Astro/Torus。

2.3.2 迭代环面构造法

如上所述，TM 程序[11, 12]给出了一般轴对称系统的相空间坐标和环面坐标转换的功能，主要是由 $(\boldsymbol{J}, \boldsymbol{\theta})$ 计算 $(\boldsymbol{x}, \boldsymbol{v})$。对于环面坐标的计算，除了按照上述方法，重复地构造环面，直到得到的相空间坐标穿过给定相空间坐标，TACT 程序的其中一个方法对其进行了改进，即迭代环面构造法 (ItTC)[22, 24]。该方法是“Stäckel 作弊法 +TM 程序”的迭代算法，即用 Stäckel 作弊法计算试探的环面坐标，然后寻找环面上与其最近的相空间点，最后补上残差并迭代[22]。

给定一个初始相空间点 $(\boldsymbol{x}, \boldsymbol{v})$，先用 Stäckel 作弊法计算一个环面坐标 $(\boldsymbol{J}_S, \boldsymbol{\theta}_S)$；然后用环面映射得到一个以此为作用量的环面，该环面的频率为 $\Omega$，在该环面上找到一个与给定点最近的相空间点 $(\boldsymbol{x}_S, \boldsymbol{v}_S)$，即

$$\eta = |\boldsymbol{\Omega}|^2(\boldsymbol{x}-\boldsymbol{x}_S)^2 + (\boldsymbol{v}-\boldsymbol{v}_S)^2 \ , \tag{69}$$

使其最小。之后，计算所找到的这一点的环面坐标 $\boldsymbol{J}_P(\boldsymbol{x}_S, \boldsymbol{v}_S)$, $\boldsymbol{\theta}_P(\boldsymbol{x}_S, \boldsymbol{v}_S)$，得到 $\Delta\boldsymbol{J} = \boldsymbol{J}_P - \boldsymbol{J}_S$ 作为误差，假设这个误差变化很慢，于是将其作为残差补回去，得到一个更好的作用量 $\boldsymbol{J}_2 = \boldsymbol{J}_S + \Delta\boldsymbol{J}$。

可以再次构造一个以 $\boldsymbol{J}_2$ 为作用量的环面，然后继续从中找到一个与给定初始点最近的相空间点,……，如此往复迭代，得到更精确的作用量，这个过程是收敛的。

2.3.3 轨道积分拟合生成函数法

这里介绍 Sanders 和 Binney[21, 24]给出的轨道积分拟合生成函数法。该方法过程为：轨道时序采样 $(\boldsymbol{x}, \boldsymbol{v})(t_i)$ → 计算玩具环面坐标 $(\boldsymbol{J}^{\rm toy}, \boldsymbol{\theta}^{\rm toy})(t_i)$ → 解生成函数系数矩阵。它和 TM 程序的区别是，该方法将环面映射思想应用于一段运动过程 (即一段时间序列)，计算出在时间 $t$ 上沿轨道采样的一系列相空间坐标的作用量、角变量和角频率，即构造生成函数是基于轨道积分，而不是 TM 所基于的某个时间点的哈密顿量偏差 $\delta H$ 的最小二乘。

具体步骤为，给定一个真实相空间点作为初始点，首先计算一个时间 $N_T T$ 的轨道积分 (画轨道)，并采集 $N_{\rm samp}$ 个相空间采样点，其中 $T$ 是具有相同能量的大圆轨道的周期；然后在每个相空间点，解析计算其玩具势中的 $(\boldsymbol{\theta}^{\rm toy}, \boldsymbol{J}^{\rm toy})$；之后将这些玩具环面坐标点代入到式 (66) 和 (67) 来求解整个方程组的未知量 (目标作用量 $\boldsymbol{J}$，生成函数傅里叶系数 $S_{\boldsymbol{n}}$ 及其导数 $\partial S_{\boldsymbol{n}}/\partial \boldsymbol{J}$，以及目标角频率 $\boldsymbol{\Omega}$ 和常数 $\boldsymbol{\theta}(0)$)。

(1) 轨道积分与轨道分类

进行轨道积分。在没有转动的情况下，一般三轴势允许两类基本的非共振轨道：环轨道和盒轨道[1, 21]。同一真实势，不同的初始条件可能导致不同类型的轨道，于是 TACT 为不同类型的轨道匹配相应的玩具势，与给定轨道类的环面相同的几何结构。

(2) 玩具势的拟合

对于环轨道，玩具势为等时势：

$$\varPhi_{\rm iso} = \frac{-GM}{b+\sqrt{b^2+r^2}} \ , \tag{70}$$

对于盒轨道，玩具势为谐振子势：

$$\Phi_{\rm har}=\frac{1}{2}\sum_{i=1}^{3}\omega_i^2x_i^2\ . \tag{71}$$

其判定依据为轨道类型，即看 $L_x$ 和 $L_z$ 的方向变化。在用玩具引力势之前，需要在轨道上采集样点，拟合出相应玩具引力势的质量和标长参数。

(3) 生成函数

在每个相空间样点 $t_i$ 计算相应的玩具作用量-角变量 $(\boldsymbol{\theta}^{\rm toy}(t_i), \boldsymbol{J}^{\rm toy}(t_i))$，并产生相应的式 (66)，其中各 $\boldsymbol{J}$ 和 $S_{\boldsymbol{n}}$ 未知。Sanders 和 Binney 轨道积分拟合生成函数法的实现技巧为："我们不能精确地解这些方程，因为处理的方程有无穷多个未知数。我们可以在每个方程的右边只包含有限个项，则方程右边和左边不能精确一致，因此可以采取的正确方法是：对每个方程的左右边的差作平方，然后对所有方程的差平方求和(即轨道积分)，并最小化这个总的平方和"[21]。其总平方和的表达式是：

$$E_1=\sum_i\sum_k\left(J_k^{\rm toy}(t_i)-J_k-2\sum_{\boldsymbol{n}\in\boldsymbol{N}}n_kS_{\boldsymbol{n}}(\boldsymbol{J})\cos\boldsymbol{n}\cdot\theta(t_i)\right)^2\ . \tag{72}$$

其中，$k=1,2,3$ 代表坐标维数，$\boldsymbol{N}$ 限制为有限维向量 (有限精度)，$|\boldsymbol{n}|\leqslant N_{\max}$，$N_{\max}\approx6$；$(t_i)$ 代表样点，最外层的求和 (关于指标 $i$) 是遍历所有采样点，此即"轨道积分"的含义。更多关于样点个数的选取原则请详见参考文献 [21] 中的 2.2 和 2.3 节。

下面最小化 $E_1$，即令 $E_1(J_k', S_{\boldsymbol{m}})$ 的偏导为 0，

$$\begin{cases}0=\dfrac{\partial E}{\partial J_k}=-2\displaystyle\sum_i\left(J_k^{\rm toy}(t_i)-J_k-2\sum_{\boldsymbol{n}\in\boldsymbol{N}}n_kS_{\boldsymbol{n}}(\boldsymbol{J}),\cos(\boldsymbol{n}\cdot\boldsymbol{\theta}^{\rm toy}(t_i))\right)\\ 0=\dfrac{\partial E}{\partial S_{\boldsymbol{m}}}=-2\displaystyle\sum_i\sum_k2m_k\cos\left(\boldsymbol{m}\cdot\boldsymbol{\theta}(t_i)\right)\\ \qquad\qquad\left(J_k^{\rm toy}(t_i)-J_k-2\displaystyle\sum_{\boldsymbol{n}\in\boldsymbol{N}}n_kS_{\boldsymbol{n}}(\boldsymbol{J})\cos\left(\boldsymbol{n}\cdot\boldsymbol{\theta}^{\rm toy}(t_i)\right)\right)\end{cases}\ . \tag{73}$$

这些方程可转化为一个矩阵方程组

$$\boldsymbol{A}_{\boldsymbol{J}^{\rm toy}}\cdot\boldsymbol{x}_{\boldsymbol{J}^{\rm toy}}=\boldsymbol{b}_{\boldsymbol{J}^{\rm toy}}\ , \tag{74}$$

其中，

$$\begin{aligned}&\boldsymbol{A}_{\boldsymbol{J}^{\rm toy}}\equiv\sum_i\begin{pmatrix}\boldsymbol{I}_3 & \boldsymbol{c}^{\rm T}(t_i)\\ \boldsymbol{c}(t_i) & \boldsymbol{c}(t_i)\cdot\boldsymbol{c}^{\rm T}(t_i)\end{pmatrix}\ ,\quad \boldsymbol{x}_{\boldsymbol{J}^{\rm toy}}\equiv(\boldsymbol{J},S_{\boldsymbol{n}})\ ,\\ &\boldsymbol{b}_{\boldsymbol{J}^{\rm toy}}\equiv\sum_i(J(t_i),\boldsymbol{c}_{\boldsymbol{n}}(t_i)\cdot\boldsymbol{J}^{\rm toy}(t_i)\ ,\quad \boldsymbol{c}_{\boldsymbol{n}k}(t_i)\equiv2n_k\cos(\boldsymbol{n}\cdot\theta(t_i))\ .\end{aligned} \tag{75}$$

其中，$\boldsymbol{c}_{\boldsymbol{n}k}(t_i)$ 为 $N$ 乘 3 的矩阵，$\boldsymbol{x}_{\boldsymbol{J}^{\rm toy}}$ 和 $\boldsymbol{b}_{\boldsymbol{J}^{\rm toy}}$ 为 $(N+3)$ 维的列向量，$\boldsymbol{I}_3$ 为单位矩阵，$\boldsymbol{A}_{\boldsymbol{J}^{\rm toy}}$ 为对称矩阵。最终求解矩阵方程得到目标作用量 $\boldsymbol{J}$ 和生成函数傅里叶系数 $S_{\boldsymbol{n}}$。

类似地，对于求角变量 $\boldsymbol{\theta}'$，则最小化

$$E_2 = \sum_i \sum_k \left( \theta'_k(0) + \Omega'_k t_i - \theta_k(t_i) - 2 \sum_{n \in \boldsymbol{N}} \frac{\partial S_{\boldsymbol{n}}(\boldsymbol{J})}{\partial J'_k} \sin \boldsymbol{n} \cdot \theta \right)^2 , \tag{76}$$

得到角频率 $\boldsymbol{\Omega}'$ 以及 $\boldsymbol{\theta}(0)$ 和 $\partial S_{\boldsymbol{n}}/\partial \boldsymbol{J}$。

此外，对式(66)中的玩具作用量$\boldsymbol{J}^{\rm toy}$ 在玩具角变量$\boldsymbol{\theta}^{\rm toy}$空间作平均，得到目标作用量

$$\boldsymbol{J} = \int \frac{{\rm d}^3 \boldsymbol{\theta}^{\rm toy}}{(2\pi)^3} \boldsymbol{J}^{\rm toy} - 2 \sum_{\boldsymbol{n} \in \boldsymbol{N}} \boldsymbol{n} S_{\boldsymbol{n}}(\boldsymbol{J}) \sin \boldsymbol{n} \cdot \boldsymbol{\theta}^{\rm toy} = \int \frac{{\rm d}^3 \boldsymbol{\theta}^{\rm toy}}{(2\pi)^3} J , \tag{77}$$

便是一种更简单的方法——平均生成函数法 (average generation function, AvGF)[24, 29]。

### 2.4　TACT 所提供的方法的效果对比

我们运行 TACT 程序[23, 24]用几种方法分别估算出作用量，如图 2 所示。我们所使用的引力势为某种轴对称引力势，即 McMillan[57] 拟合的最佳银河系质量分布模型，分为薄厚两个恒星盘、HI 和 $H_2$ 两个气体盘以及核球和暗物质晕；各个星系成分的模型输入参数(适合 TACT 程序的格式)，已在作者的 github 主页① 给出。我们采用的恒星轨道为典型的银河系薄盘轨道，初始条件为 $(\boldsymbol{x}, \boldsymbol{v})$ = (8.178, 0.1, 0.225; 20.13, 187.01, 14.95)，单位为 kpc 和 km/s。图中的几种方法分别是：Fudge v1 和 Fudge v2 为 Stäckel Fudge 作弊法，其中，Fudge v1 用式 (48) 估算焦距，Fudge v2 基于壳轨道 (shell orbits) 来估算焦距；Fit 对应 Stäckel 势拟合法，ItTC 对应环面迭代法，O2GF 对应轨道积分拟合生成函数法，AvGF 为平均生成函数法；CAA 和 SAA 分别为柱绝热近似和球绝热近似方法，下一节作简介。轨道中的每个点都独立计算一次作用量，因而作用量的数值随时间有所波动。

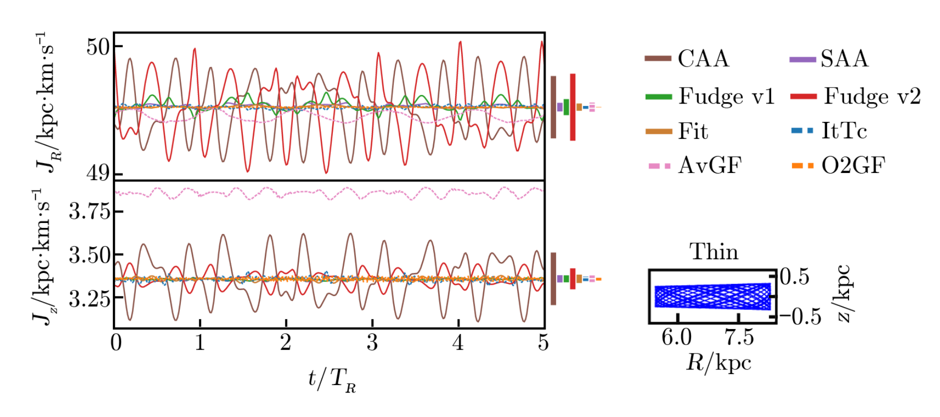



图 2　作用量数值计算的效果对比

Sanders 和 Binney[24] 对几种方法的在轨道类型、精度和时间等方面的性能做了更详细的对比(见其中的图 2 到图 6)。总体上看，收敛方法比非收敛方法有更高的精度，但耗时也高 1 ~ 3 个量级。其中，轨道积分拟合生成函数法精度最高，耗时较长；Stäckel 作弊法速

①https://github.com/PaulMcMillan-Astro/GalPot/blob/master/pot/PJM17_best.Tpot。

度很快，是大量计算的最经济方法，但精度不如收敛性方法高，特别是对于盒轨道的误差大。另外，对于轴对称势经过环面迭代后 (即 ItTC 法，目前 TM 程序只提供了轴对称的环面迭代)，Stäckel 作弊法的精度会提高，而耗时也变长。此外，TACT 程序包中，轨道积分拟合生成函数法和 Stäckel 作弊法还能用于三轴势。

### 2.5　其他一些求运动积分的方法或程序

前面两节我们主要介绍了 Binney 及其合作者们近年来的工作。接下来我们介绍其他一些基于 Stäckel 势的估算运动积分或作用量的方法。

#### 2.5.1　“近似的运动积分”解析式

Bienaymé[54]给出了一种通过椭球坐标轴上的代表点来计算所有位置处的引力势的方法，因此一般引力势可以拟合为 Stäckel 形式。然后，就可以根据 Stäckel 势的运动积分解析表达式，给出这个一般引力势下的“近似的运动积分”解析式 (又称为“准运动积分”，quasi integral of motion[53]；特别重要的是“quasi-$I_3$”的表达式)。该方法适应接近 Stäckel 形式的势。下面介绍这种方法。

由式 (21)，将 $P_\tau^2$ 代入，可写为：

$$\begin{aligned}\varPhi = &-4\frac{\zeta(\lambda)(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)}{(\lambda-\mu)(\lambda-\nu)} - 4\frac{\eta(\mu)(\mu+\alpha)(\mu+\beta)(\mu+\gamma)}{(\mu-\nu)(\mu-\lambda)} - \\ &4\frac{\kappa(\nu)(\nu+\alpha)(\nu+\beta)(\nu+\gamma)}{(\nu-\lambda)(\nu-\mu)} \quad .\end{aligned} \tag{78}$$

若 $\varPhi$ 在 $(-\alpha,\ -\beta,\ -\gamma)$ 连续，应有 $\varPhi(-\alpha,\ -\beta,\ -\gamma)=0$。让势两边取 $(\lambda,\ -\beta,\ -\gamma)$ 位置的值，则有

$$\zeta(\lambda) = -\frac{\varPhi(\lambda,-\beta,-\gamma)}{4(\lambda+\alpha)} \quad , \tag{79}$$

同理有

$$\eta(\mu) = -\frac{\varPhi(-\alpha,\mu,-\gamma)}{4(\mu+\beta)} \quad , \quad \kappa(\nu) = -\frac{\varPhi(-\alpha,-\beta,\nu)}{4(\nu+\gamma)} \quad . \tag{80}$$

将上面的分离函数代回到式 (78)，可得

$$\begin{aligned}\varPhi(\lambda,\mu,\nu) = &\frac{(\lambda+\beta)(\lambda+\gamma)}{(\lambda-\mu)(\lambda-\nu)}\varPhi(\lambda,-\beta,-\gamma) + \frac{(\mu+\alpha)(\mu+\gamma)}{(\mu-\nu)(\mu-\lambda)}\varPhi(-\alpha,\mu,-\gamma) + \\ &\frac{(\nu+\alpha)(\nu+\beta)}{(\nu-\lambda)(\nu-\mu)}\varPhi(-\alpha,-\beta,\nu) \quad .\end{aligned} \tag{81}$$

这样就能分别用一个“椭球坐标轴”(如 $(\lambda,\ -\beta,\ -\gamma)$ 中后两个坐标固定) 上的势值来表示任意点的引力势。

采取类似的手法，这个引力势下的运动积分可以解析性地表达出来。运动积分 $J$, $K$ 的表达式为

$$\begin{cases} J = 1/2L^2 - 1/2(\beta+\gamma)v_x^2 - 1/2(\gamma+\alpha)v_y^2 - 1/2(\alpha+\beta)v_z^2 - \\ \quad (\mu+\nu)\dfrac{\zeta(\lambda)}{P_\lambda^2} - (\nu+\lambda)\dfrac{\eta(\mu)}{P_\mu^2} - (\lambda+\mu)\dfrac{\kappa(\nu)}{P_\nu^2} \\ K = 1/2\beta\gamma v_x^2 + 1/2\gamma\alpha v_y^2 + 1/2\alpha\beta v_z^2 - 1/2\alpha L_x^2 - 1/2\beta L_y^2 - 1/2\gamma L_z^2 - \\ \quad \mu\nu\dfrac{\zeta(\lambda)}{P_\lambda^2} - \nu\lambda\dfrac{\eta(\mu)}{P_\mu^2} - \lambda\mu\dfrac{\kappa(\nu)}{P_\nu^2} \end{cases} , \qquad (82)$$

相关推导见参考文献 [2] 中的式 (9) 和 (14)。将式 (79) 和 (80) 代入，可计算得到与式 (78) 同样方式的运动积分 $J$, $K$ 的表达式。

类似地，由式 (15)，$I_2$, $I_3$ (第一个运动积分是能量 $E$) 的表达式可写为：

$$\begin{cases} I_2 = \psi(\lambda,\mu,\nu) + 1/2(\alpha-\beta)v_x^2 + 1/2\dfrac{(\alpha-\beta)}{(\alpha-\gamma)}L_y^2 + 1/2L_z^2 \\ I_3 = \xi(\lambda,\mu,\nu) + 1/2(\gamma-\beta)v_z^2 + 1/2\dfrac{(\gamma-\beta)}{(\gamma-\alpha)}L_y^2 + 1/2L_x^2 \end{cases} , \qquad (83)$$

即参考文献 [54] 中的式 (9) 和 (10)。文献 [54] 推导了上式中的椭球坐标的函数项，得到

$$\begin{cases} \psi = \dfrac{(\mu+\alpha)(\nu+\alpha)}{(\alpha-\gamma)}\Phi(\lambda,\mu,\nu) , \\ \xi = \dfrac{(\mu+\gamma)(\lambda+\gamma)}{(\gamma-\alpha)}\Phi(\lambda,\mu,\nu) , \end{cases} \qquad (84)$$

其中，$\Phi(\lambda,\mu,\nu)$ 即为式 (78) 中的形式，其简化形式见文献 [54] 中的式 (16) 和 (17)。

该方法计算“每一点的”运动积分，需要当前点的坐标、速度和角动量等轨道信息。文献 [54] 在后文用高度精确的数值方法计算出轨道，并研究了这种方法所计算的运动积分的误差，最后做了 Jeans 方程的检验。

### 2.5.2 绝热近似

对于轴对称势，文献 [24] 还介绍了用近似公式来估算作用量的方法。通过假设在柱坐标系或球坐标系中引力势是可分离的，可求出这些作用量。

(1) 柱绝热近似[45]。这种观点认为，盘轨道恒星的垂向 (即 $z$ 方向) 频率 $\Omega_z$ 比其径向频率 $\Omega_r$ 大得多，因此当恒星径向振荡时，主导其垂向振荡的势可被认为变化缓慢，其结果就是 $J_z$ 是绝热不变的[24]。用这些假设推导出作用量的近似公式

$$\begin{cases} J_z = \dfrac{2}{\pi}\displaystyle\int_0^{z_{\max}} \mathrm{d}z v_z \\ J_r = \dfrac{1}{\pi}\displaystyle\int_{R_{\rm peri}}^{R_{\rm apo}} \mathrm{d}R v_R \end{cases} , \qquad (85)$$

以及计算积分上下限时需要用到的垂直能量 (即垂直动能加垂直势能，其中前者是由于速度的平方按正交坐标方向分为三个分量平方的相加；后者是保守力按照某方向的分解所得的

势能) 和有效势：

$$\begin{cases} E_z = \frac{1}{2}v_z^2 + \Phi(R,z) - \Phi(R,0) \\ \Phi_R = \Phi(R,0) + \frac{J_\phi^2}{2R^2} + E_z(J_z,R) - E_z(J_z,R_c) \end{cases} , \tag{86}$$

其中，$R_c$ 是导心半径 (可参见有关“本轮近似”的文献)。

(2) 球绝热近似假设沿长球坐标的球面存在绝热不变振荡。作用量的近似公式为：

$$\begin{cases} J_z = \frac{2}{\pi}\int_{c^2}^{\nu_+} \mathrm{d}\nu p_\nu = \frac{2}{\pi}\int_{c^2}^{\nu_+} \mathrm{d}\nu\sqrt{2P_\nu^2(\lambda,\nu)}\sqrt{E_\nu - \Phi_\nu(\nu)} \\ J_r = \frac{1}{\pi}\int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda p\lambda = \frac{1}{\pi}\int_{\lambda_-}^{\lambda_+} \mathrm{d}\lambda\sqrt{2P_\lambda^2(\lambda,c^2)}\sqrt{E_\lambda - \Phi_\lambda(\lambda)} \end{cases} , \tag{87}$$

其中，

$$\begin{cases} \Phi_\nu(\nu) = \frac{J_\phi^2}{2R^2(\lambda,\nu)} - \frac{J_\phi^2}{2R^2(\lambda,c^2)} + \Phi(\lambda,\nu) - \Phi(\lambda,c^2) \\ \Phi_\lambda(\lambda) = \Phi(\lambda,c^2) + \frac{J_\phi^2}{2R^2(\lambda,c^2)} + E_\nu(R=\sqrt{\lambda-a^2},J_\phi,J_z) \\ E_\tau = \frac{p_\tau^2}{2P_\tau^2} + \Phi_\tau(\tau)\ ,\ \ \tau=\nu,\lambda \end{cases} . \tag{88}$$

具体计算可见参考文献 [24]。

2.5.3　一些相关程序

除了 Sanders 和 Binney[24] 的 TACT 程序包以及其环面迭代功能所依赖的 TM 程序包，星系动力学还有一些程序可供使用，这些程序现在都可以公开获取。主要程序如下所述。

(1) Gadget，Gadget① 以及其后续发展 Arepo②程序，是“引力 $N$ 体+流体力学”模拟程序。在星系动力学领域，我们可以用它来做无碰撞 $N$ 体模拟 (特别是当我们考察暗物质粒子构成的暗物质晕的动力学时)。

(2) Galpy，Galpy (a Python library for galactic dynamics)③ 是 Bovy[30] 提供的 Python 程序包，能用来进行数值轨道积分、作用量-角变量、分布函数等相关计算。

(3) AGAMA，AGAMA (action-based galaxy modelling architecture)④ 是 Vasiliev[31] 提供的星系动力学程序，能提供任意解析密度剖面或 $N$ 体模型的引力势的计算、轨道积分和分析、位置-速度与作用-角度变量之间的转换、分布函数的计算以及迭代构造自洽多组分星系模型等功能。

①https://wwwmpa.mpa-garching.mpg.de/galform/gadget/index.shtml

②https://arepo-code.org

③https://github.com/jobovy/galpy

④https://github.com/GalacticDynamics-Oxford/Agama

2.5.4 深度学习在作用量计算中的应用

在这一小节中，我们介绍最新的一个研究动态。Ibata 等人[32]将人工智能 (AI) 中的深度学习方法应用到作用量计算上。该文用 AI 算法来实现在静态引力势下从粒子在相空间的“位置-速度”到“作用量-角变量”的映射，并且得到加速度场。该算法的核心原理与前述“生成函数法”类似，先从数据点中拟合玩具势；然后通过从玩具环面 ($\boldsymbol{\theta}$, $\boldsymbol{J}$) 到目标环面 ($\boldsymbol{\theta}'$, $\boldsymbol{J}'$) 的正则变换表达式，迭代出最合适的生成函数 $G$；之后是正则点变换的生成函数 $P$，最终得到所求的目标环面 ($\boldsymbol{\theta}''$, $\boldsymbol{J}''$)。其中两个正则变换所用网络的形式为：

$$G = \mathrm{net}_G(\cos\boldsymbol{\theta},\ \sin\boldsymbol{\theta},\ \boldsymbol{J}')\ , \quad P = \mathrm{net}_P(\cos\boldsymbol{\theta},\ \sin\boldsymbol{\theta})\ . \tag{89}$$

该算法的程序为 ACTIONFINDER，基于 Python 的 Pytorch 所写。据文献 [32] 介绍，经测试 ACTIONFINDER 得到的作用量精度要比 Stäckel 作弊法要高，而且它还有如下优点：相比于早期的 Torus Mapper 对初值要求更不敏感；不需要向它提供系统内轨道的公平样本，几乎可覆盖任何感兴趣区域的样本；不需要预先知道所研究系统的哈密顿量或引力势；此外，该算法所用的深层神经网络，相比前述“生成函数法”线性分解成的傅里叶系数，能够生成更加灵活的非线性函数，并且这种灵活性可能使其更容易拟合更一般的动力系统。看来，将来经过一定的发展完善后，深度学习或其他AI技术用来计算作用量，适用性会比传统方法更广泛、性能更佳，应该有一定的发展前景，未来或成为分析更复杂体系的另一种尝试。

## 3 应用

在星系研究中，作用量的一个重要应用在分布函数方面，即使用以作用量作为自变量的分布函数，为星系建模[15]。一般来说，我们难以直接获得分布函数的形式。在建模以及数据拟合的实际工作中，人们通常是为不同的星系组分 (如恒星盘、核球、暗物质晕等) 采用某种先验的分布函数形式，然后根据数据拟合来确定其中的具体参数的取值。

大约 10 年前，基于分布函数的建模方法实际上很难应用 (见 Read[61] 和 Binney[15] 的综述)，星系的建模工作往往是基于 Jeans 方程，即对分布函数所满足的无碰撞玻尔兹曼方程取矩 (即所谓“矩方法”)。假定星系处于稳态 (即 $\partial f/\partial t = 0$)，从无碰撞玻尔兹曼方程出发，分别乘上速度分量并且积分掉速度空间 (即取矩)，即得到 Jeans 方程。对于实际的银河系建模工作，研究者还会假设轴对称 (即在柱极坐标下 $\partial f/\partial\phi = 0$, $\partial f/\partial v_\phi = 0$)，那么 Jeans 方程的表达式为 (柱极坐标)：

$$\begin{cases} \dfrac{\nu\sigma_R^2}{\partial R} + \dfrac{\partial\nu\,\sigma_{Rz}}{\partial z} + \nu\left(\dfrac{\sigma_R^2 - \overline{v_\phi^2}}{R} + \dfrac{\partial\Phi}{\partial R}\right) = 0\ , \quad R\ 方向 \\ \dfrac{1}{R^2}\dfrac{\partial(R^2\nu\,\overline{v_R v_\phi})}{\partial R} + \dfrac{\partial(\nu\,\overline{v_z v_\phi})}{\partial z} = 0\ , \quad \phi\ 方向 \\ \dfrac{1}{R}\dfrac{\partial(R\nu\,\sigma_{Rz}^2)}{\partial R} + \dfrac{\partial(\nu\sigma_z^2)}{\partial z} + \nu\dfrac{\partial\Phi}{\partial z} = 0\ , \quad z\ 方向 \end{cases} \tag{90}$$

其中，$\nu$ 为示踪者恒星的数密度，为分布函数的 0 阶矩；平均速度 $\overline{v_i}$ 为 1 阶矩；速度弥散张量为

$$\sigma_{ij} = \frac{1}{\nu}\int \mathrm{d}^3\boldsymbol{v}(v_i - \overline{v_i})(v_j - \overline{v_j})f(\boldsymbol{x},\boldsymbol{v}) \quad , \tag{91}$$

为 2 阶矩。在观测上，速度弥散可由某一个位置处诸恒星的速度方差来计算。

正如 Read 所总结[61]，Jeans 方程和矩方法的优点是：与其他方法相比非常快，可以探索较大的参数空间；并且不需要假设分布函数的形式，因为只是限制了它的矩。缺点是：必须对数据进行整理以计算速度的 2 阶矩 (但实际上恒星的速度数据往往不是高斯分布，因此不能完全为二阶矩所刻画)；分布函数的形式不能使用 (因此未能使用恒星完全的分布信息)；Jeans 方程组一般情况下不能闭合 (即不能求解)，在某些情况下可能得到一个实际分布函数不存在的解。

如果这时候我们进一步假设该星系中的恒星运动只遵守两个运动积分 (即文献中常表述的 $f(E,L_z)$)，那么 $\sigma_R = \sigma_z$，而且 $\sigma_{Rz} = 0$。此时，$z$ 方向的 Jeans 方程，就仅与各量的 $z$ 方向的梯度有关，与 $R$ 方向的运动解除耦合了，即：

$$\frac{\partial(\nu\sigma_z^2)}{\partial z} = -\nu\frac{\partial\Phi}{\partial z} \quad . \tag{92}$$

在给定 $\nu, \Phi$ 之后，我们可以得到 $\sigma_R$ (或 $\sigma_z$) 的表达式以及 $\overline{v_\phi^2}$ 的表达式，也就是说 Jeans 方程闭合了[8]。

在更早之前，就连上述 2 个运动积分情形下的 Jeans 方程都难以得到银河系观测数据 (恒星的密度分布数据以及运动学数据) 的支持，研究者把引力势的表达式 (即泊松方程) 也“一维化”了，简化为：

$$4\pi G\rho = \frac{\partial^2\Phi}{\partial z^2} \quad . \tag{93}$$

式 (92) 和 (93) 构成了所谓的“一维近似”[61]。由于 $-\partial\Phi/\partial z$ 即垂向力 $K_z$ (引力的 $z$ 方向分量)，早期文献中又把这种一维近似称作“$K_z$ force”方法[39]。如果用更方便的观测量恒星面密度 ($\Sigma$) 来代替密度 $\rho$，那么泊松方程可以表示为 $K_z = 2\pi G\Sigma$。

在下面的几个小节中，本文介绍近年来有代表性的建模工作。这几个工作都是基于 3 个运动积分 (或作用量) 的分布函数来为银河系建模。这些工作中，对于观测到的恒星盘成分，一般是使用基于作用量的准等温球分布函数模型，或者基于运动积分的准等温球分布函数。对于暗物质晕，这些工作使用质量分布模型 (如 NFW 模型) 而不是分布函数模型。后来，Piffl 等人[51]及 Binney 和 Piffl[49]开始尝试使用分布函数模型来直接为暗物质晕建模并拟合银河系数据；这项暗物质晕建模工作由于数据有限以及软件上的问题，目前的效果不太理想[15]。在本节的末尾，我们还介绍前人提出的针对恒星盘的另一种分布函数模型。

前人的这些工作，主要是基于 RAVE 和 SDSS 的银河系巡天数据。现在，在 Gaia 的天测和测光数据以及大规模光谱巡天 (例如中国的 LAMOST) 的基础上，我们预计这些基于分布函数建模的方法将大放异彩。

### 3.1 基于 $f(J_R, L_z, J_z)$ 的银河系动力学建模

#### 3.1.1 Bovy 和 Rix 的模型

Bovy 和 Rix[44]用 SEGUE 的 G 矮星样本 (范围为 $5 < R_{\rm GC} < 12\,{\rm kpc}$, $0.3 \lesssim |z| \lesssim 3\,{\rm kpc}$，其中，$R_{\rm GC}$ 为到银河系中心的径向距离，$|z|$ 为到银河系中心的垂向距离) 进行建模，来推断恒星和暗物质质量分布。将数据按金属丰度分为 43 个单丰度亚群 (mono-abundance population, MAP)，对每个 MAP 的相空间数据独立地设置一个参数化的银河系引力势和一个基于作用量的准等温的恒星分布函数；并且假设每个 MAP 的数据在不同的半径约束垂向力为 $K_z$。用这些结果，研究了在 $4.5\,{\rm kpc} < R_{\rm GC} < 9\,{\rm kpc}$ 范围，银河系中平面 1.1 kpc 内的总面密度 $\Sigma_{1.1}(R) = \int_{-1.1\,{\rm kpc}}^{1.1\,{\rm kpc}} {\rm d}z\rho(R,z)$ 和垂向力 $K_z$，以及暗物质晕的动力学贡献等内容。

(1) 恒星的分布函数模型

准等温分布模型 (quasi-distribution function, qDF)[45] 是一种最常见的恒星分布模型。将示踪星的分布函数建模为一个或多个等温分量；根据定义，等温分量具有恒定的速度色散，与 $z$ 无关，从而速度分布必须是高斯分布[39]。Bovy 和 Rix[44]文中的每个 MAP 都是由准等温分布拟合的恒星种群；分布函数的自由参数，与示踪剂群体的径向密度分布，径向和垂向速度色散，以及它们如何随 $R$ 变化有关。Bovy 和 Rix[44]所用的基于作用量的恒星分布模型来自文献 [46]：

$$
{\rm qDF}(\boldsymbol{x},\boldsymbol{v}) = f_{\sigma_R}(J_R, L_z) \times \frac{\nu}{2\pi\sigma_z^2} \times \exp\left(-\frac{\nu J_z}{\sigma_z^2(R_c)}\right) \quad , \tag{94}
$$

其中，$f_{\sigma_R}(J_R, L_z) = \dfrac{\Omega n(R_c)}{\pi\kappa\sigma_R^2(R_c)}|_{R_c} \times [1+\tanh(L_z/L_0)] \times \exp\left(-\dfrac{\kappa J_R}{\sigma_R^2(R_c)}\right)$，$\kappa$，$\Omega$，$\nu$ 是本轮频率、圆周频率和垂向频率，$R_c$ 是具有角动量 $L_z$ 的圆形轨道的半径。并且假设示踪者密度径向轮廓、径向和垂向速度弥散是指数形式的：

$$
\begin{cases}
n(R_c) \propto \exp(-R_c/h_R) \\
\sigma_R(R_c) = \sigma_{R,0}\exp(-(R_c - R_0)/h_{\sigma_R}) \\
\sigma_z(R_c) = \sigma_{z,0}\exp(-(R_c - R_0)/h_{\sigma_z})
\end{cases} \quad . \tag{95}
$$

分布函数中的作用量，采用文献 [13] 的 Stäckel 作弊法来计算。

(2) 银河系引力势模型

为计算轨道作用量，Bovy 和 Rix[44]提出一个完整的三维模型来计算银河系的势：一个先验有许多自由参数的“核球+恒星盘+气体盘+暗物质晕”模型。

指数盘密度为双幂律盘模型：

$$
\rho_{\rm disk}(R,z) = \rho_d \exp(-R/R_d - |Z|/z_h) \ , \tag{96}
$$

引力势通过该密度的泊松方程的解析解得到。

核球模型为：

$$
\rho_{\rm bulge}(r) = \frac{\rho_0}{(r/a)^{\alpha}(1+r/a)^(\beta-\alpha)} \ , \tag{97}
$$

其中，$\alpha=1$, $\beta=4$ 时，称为 Hernquist 模型；$\alpha=2$, $\beta=4$ 时，称为 Jaffe 模型。暗物质晕为幂律模型：

$$\rho_{\rm DM}(R,z)\propto 1/r^{\alpha}\ , \tag{98}$$

它们的共同引力势为总引力势。

在 Bovy 和 Rix[44] 对单个 MAP 的 qDF 拟合中，除了恒星盘的标长 $R_d$ 和晕对盘在 $R_0$ 处的径向力的贡献以外，其他基本参数 (恒星盘的标高 $z_h$，以及 $R_0$ 附近圆周速度 $v_c(R_0)$ 和旋转曲线的“平坦度” $\mathrm{d}\lg v_C(R_0)/\mathrm{d}\lg(R)$) 都被固定。

(3) 拟合手续

Bovy 和 Rix[44] 所使用的拟合方法为最大似然法，其拟合步骤如下，流程示意图见图 3。

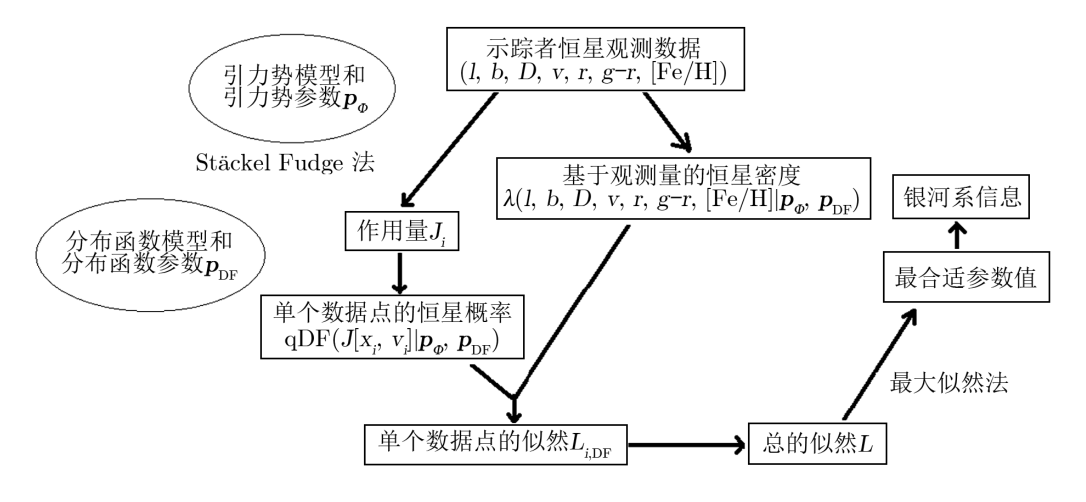



图 3　Bovy 和 Rix 拟合流程的示意图[44]

观测量下的分布为：

$$\begin{aligned}\lambda(l,b,D,\boldsymbol{v},r,g-r,[\mathrm{Fe/H}]) =&\rho(r,g-r,[\mathrm{Fe/H}]||R,z,\phi)\times \mathrm{qDF}(\boldsymbol{x},\boldsymbol{v})\times\\ &|J(R,z,\rho;l,b,D)|\times S(\mathrm{plate},r,g-r)\ ,\end{aligned} \tag{99}$$

其中，$D$ 为观测量距离，$l$ 为银经，$b$ 为银纬，$g-r$ 为颜色，[Fe/H] 为观测量金属丰度，$\boldsymbol{v}$ 为速度，由观测得到，$(R,z,\Phi)$ 为银心柱坐标系，$|J|$ 为坐标变换的雅克比，$S$ 为选择函数。

归一化后，每个数据点的似然为：

$$\begin{aligned}\ln\mathscr{L}_{i,\rm DF}=&\ln\mathrm{qDF}(J(\boldsymbol{x}_i,\boldsymbol{v}_i)|\boldsymbol{p}_\Phi,\boldsymbol{p}_{\rm DF})-\\ &\ln\int \mathrm{d}l\mathrm{d}b\mathrm{d}D\mathrm{d}\boldsymbol{v}\mathrm{d}r\mathrm{d}(g-r)\mathrm{d}[\mathrm{Fe/H}]\lambda(l,b,D,\boldsymbol{v},r,g-r,[\mathrm{Fe/H}]|\boldsymbol{p}_\Phi,\boldsymbol{p}_{\rm DF})\ .\end{aligned} \tag{100}$$

加上参数条件是表示设参数 $\boldsymbol{p}_\Phi$, $\boldsymbol{p}_{\rm DF}$ 取某一组值的条件，最大似然就是在参数取值空间中找到似然最大时的参数值。作用量的计算会用到总引力势 $\Phi(R,z)$，这些引力势的参数实际

是待定的基准势；正是在给定分布函数模型和引力势模型后，用最大似然法依靠数据从参数空间找到最合适的参数后，才能得到最合适的模型。因此，引力势的参数，正是通过数据对作用量的依赖，输入到观测到的密度当中[44]。

数据点是独立的，考虑异常数据点 (outliers) 后，总的似然为：

$$\ln \mathscr{L} = \sum_i \ln[(1 - p_{\rm out}) \ln \mathscr{L}_{i,\rm DF} + p_{\rm out} \ln \mathscr{L}_{i,\rm out}] \ . \tag{101}$$

最后根据最大似然法估计，使这个总似然最大，即找到所求参数 $\boldsymbol{p}_\Phi$，$\boldsymbol{p}_{\rm DF}$ 的最合适值。

(4) 结果

利用这些模型的所有参数的最佳结果，Bovy 和 Rix[44]得到了星系盘的径向标长 $R_d = (2.15 \pm 0.14)\,\rm kpc$，与通过光度法推断的恒星质量的空间分布一致。

来自不同的 MAP 数据，以不同的半径约束 $\Sigma_{1.1}(R)$；因而 Bovy 和 Rix[44]用这 43 个 MAP 数据的平均半径和 $\Sigma_{1.1}$ 的平均值的关系，研究了 $\Sigma_{1.1}$ 随半径的变化，这些 MAP 的结果也一致[44]。类似地，还得到了 1.1 kpc 的垂向力 $K_{z,1.1}(R)$。使用从文献 [41] 距离标度的后验分布函数所导出的最佳半径，将这些分别拟合为：

$$\begin{cases} \Sigma_1.1(R) = 70 M_\odot \mathrm{pc}^{-2} \exp(-(R - R_0)/(2.5\,\mathrm{kpc})) \\ K_{z,1.1}(R) = 68 M_\odot \mathrm{pc}^{-2} \exp(-(R - R_0)/(2.7\,\mathrm{kpc})) \end{cases} . \tag{102}$$

Bovy 和 Rix[44]在计算恒星盘对圆周速度的贡献、暗物质晕的幂律指数等结果方面，对完整 Jeans 方程和垂向力两种方法进行了对比 (见文献 [44] 中的图 18, 19, 20)；还展示了恒星盘和暗物质晕分别对旋转速度的局部贡献 (见文献 [44] 中的图 21)；得到的对暗物质晕的径向轮廓的限制为：在 95% 置信度下，$\rho_{DM}(r; UR0) \propto 1/r^\alpha$，$\alpha < 1.53$。

### 3.1.2 Piffl 等人的模型

与 Bovy 和 Rix[44]提出的模型类似，同样用数据限制银河系的参数，Piffl 等人[50]用距银河系平面约 1.5 kpc 内约 200 000 颗巨星的运动学，来测量太阳附近质量密度的垂向分布，以及局部暗物质晕贡献。其所用的模型的组分更多，由气体盘、薄盘、厚盘、核球和暗物质晕五个部分组成，下面介绍其质量模型和分布函数。

(1) 质量模型

盘的密度模型为：

$$\rho_1(R, z) = \frac{\Sigma_0}{2z_d} \exp\left[-\left(\frac{R}{R_d} + \frac{z}{z_d} + \frac{R_{\rm hole}}{R}\right)\right] \ , \tag{103}$$

其中，气体盘非零参数 $R_{\rm hole}$ 是为了在圆盘中创建中心腔，而其他两个盘则设置为零。

核球和暗物质晕的密度模型为：

$$\rho_2(R, z) = \frac{\rho_0}{m^\gamma (1 + m)^{(\beta - \gamma)}} \exp[-(m r_0 / r_{\rm cut})^2] \ , \tag{104}$$

其中，$m(R, z) = \sqrt{\left(\frac{R}{r_0}\right)^2 + \left(\frac{z}{q r_0}\right)^2}$，$\rho_0$ 设置密度标度，$r_0$ 为半径标度，$q$ 为等密度面的

轴比，cut 为截断；指数 $\gamma$, $\beta$ 控制径向密度轮廓的内外径向斜率；对于核球和暗物质晕，分别采用不同的轴比和截断等[50]。公式中的其他参数的意义与上一节的相同。

因此 Piffl 等人[50]的质量模型具有 8 个参数，薄盘和厚盘分别具有 3 个自由参数 ($\Sigma_0$, $R_d$, $z_d$)，而暗物质晕具有 2 个自由参数 $\rho_0, r_0$，其中薄盘和厚盘的径向标长取相同。文献 [50] 中的表 1 给出了银河系模型的固定参数取值。

(2) 分布函数模型

分布函数分为盘和晕两大部分，

$$f(J_r, J_z, L_z) = f_{\rm disk}(J_r, J_z, L_z) + F_{\rm halo} f_{\rm halo}(J_r, J_z, L_z) \ , \tag{105}$$

其中，$f_{\rm disk}$ 为盘的分布函数，$f_{\rm halo}$ 为晕的分布函数。$F_{\rm halo}$ 是晕与盘的质量之比。

盘的分布函数为：

$$f_{\rm disk}(J_r, J_z, L_z) = f_{\rm thin}(J_r, J_z, L_z) + F_{\rm thick} f_{\rm thick}(J_r, J_z, L_z) \ , \tag{106}$$

其中，$f_{\rm thin}$ 为薄盘的分布函数，$f_{\rm thick}$ 为厚盘的分布函数，$F_{\rm thick}$ 是厚盘与薄盘的质量之比。

盘的分布函数也是基于作用量的，并且通过 Stäckel 作弊法[13]估算。类似于前人的准等温模型，盘分布函数也是由准等温组分构成，

$$\begin{cases} f_{\sigma_r}(J_r, L_z) = \dfrac{\Omega\Sigma}{\pi\sigma_r^2\kappa}[1+\tanh(L_z/L_0)]\exp(-\kappa J_r/\sigma_r^2) \\ f_{\sigma_z}(J_r, L_z) = \dfrac{\nu}{2\pi\sigma_z^2}\exp(-\nu J_z/\sigma_z^2) \end{cases} , \tag{107}$$

其中，$\nu$ 仍代表垂向频率而不是椭球坐标。但是 Piffl 等人[50]的模型中，薄盘还考虑了星族年龄，其速度弥散除了 $L_z$，还依赖于年龄 $\tau$，

$$\begin{cases} \Sigma(L_z) = \Sigma_0\exp(-R_c/R_d) \\ \sigma_r(L_z,\tau) = \sigma_{r_0}\left(\dfrac{\tau+\tau_1}{\tau_m+\tau_1}\right)^\beta \exp(R_0-R_c)/R_{\sigma_r} \\ \sigma_z(L_z,\tau) = \sigma_{z_0}\left(\dfrac{\tau+\tau_1}{\tau_m+\tau_1}\right)^\beta \exp(R_0-R_c)/R_{\sigma_z} \end{cases} . \tag{108}$$

进一步假设薄盘中的恒星形成率随时间指数下降，特征时间标度为 $t_0$，于是薄盘的分布函数为：

$$f_{\rm thin}(J_r, J_z, L_z) = \frac{\int_0^{\tau_m}{\rm d}\tau {\rm e}^{\tau/t_0} f_{\sigma_r}(J_r, L_z) f_{\sigma_z}(J_z, L_z)}{({\rm e}^{\tau_m/t_0}-1)t_0} \ . \tag{109}$$

Piffl 等人[50]还在分布函数中添加了用于恒星晕的组分 $f_{\rm halo}$，以防止拟合过程使厚恒星盘变形，从而解释样本中的晕星的存在。恒星晕的分布函数为 Posti 等人[27]的幂律模型

$$f_{\rm halo}(J_r, J_z, L_z) = \left(\frac{h(L_{c_0})}{h(\boldsymbol{J})}\right)^\alpha \ , \tag{110}$$

其中，$h(\boldsymbol{J}) = J_r + \dfrac{\Omega_z}{\Omega_r}J_z + \dfrac{\Omega_\phi}{\Omega_r}|L_z|$，$h(\boldsymbol{J})$ 是 $\boldsymbol{J}$ 的一次齐次函数。

Piffl 等人[50]采用最小二乘法来建立这些模型，先给出了最佳拟合的待定参数，然后给出了各组分的密度轮廓，他们还分别给出了重子物质与暗物质晕对圆周速度和中盘面总面密度随径向的局部贡献变化。

### 3.2 基于 $f(E, L_z, I_3)$ 的银河系动力学建模

前面几个准等温模型及其变形都是基于作用量的分布函数模型。除了基于作用量的，以前也有直接基于 $E$, $L_z$, $I_3$ 三个运动积分的分布函数模型。关于第三运动积分，大体上也是以 Stäckel 引力势中的运动积分的形式为蓝本，作近似假设。下面简介其中的一些模型。

#### 3.2.1 Bienaymé 等人的模型

Bienaymé 等人[52]也用 RAVE 的红团簇星数据研究了径向力和暗物质晕的局部贡献。其假设银河系引力势为某种 Stäckel 势：

$$\varPhi(\lambda, \nu) = -\frac{h(\lambda) - h(\nu)}{\lambda - \nu} \ , \tag{111}$$

其中，$h(\tau) = GM\sqrt{\tau + q}$ 为 Kuzmin-Kutuzov 势[5, 52]类 (多了参数 $q$)。这种轴对称 Stäckel 势的第三运动积分 (见文献 [3] 的中表 1 和文献 [2] 中的式 (16)) 为：

$$I_3 = \frac{(\nu+\gamma)h(\lambda) - (\lambda+\gamma)h(\nu)}{(\gamma-\alpha)(\lambda-\nu)} - \frac{1}{2}\frac{z^2}{\gamma-\alpha}(\dot{r}^2 + (r\dot{\theta})^2) - \frac{1}{2}\left(\frac{r^2}{\gamma-\alpha} + 1\right)\dot{z}^2 + \frac{rz\dot{r}\dot{z}}{\gamma-\alpha} \ , \tag{112}$$

这里，$(r, z, \varPhi)$ 为柱坐标，$z_0 = \pm\sqrt{\gamma - \alpha}$ 为 $z$ 轴上的焦距。

恒星盘分布函数假定为一个三运动积分模型：

$$\begin{aligned} f(E, L_z, I_3) =& 2\frac{\Omega(R_c)\Sigma(L_z)}{2\pi\kappa\sigma_r^2(L_z)}\exp\left(-\frac{E - E_{\rm circ}}{\sigma_r^2}\right) \times \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_z(L_z)} \\ & \exp\left\{-\left(\frac{R_c^2(L_z)}{z_0^2} + 1\right)^{-1}\left(\frac{1}{\sigma_z^2} - \frac{1}{\sigma_r^2}\right)I_3\right\} \ . \end{aligned} \tag{113}$$

Bienaymé 等人[52]也是用 $\chi^2$ 方法拟合参数。由式 (113) 来计算矩 (从分布函数得到的速度的若干阶平均)，在观测数据上，通过拟合从三个金属样本中的观测得到矩，即数密度 $\nu(z)$ 和垂向速度弥散 $\sigma_{zz}(z)$ 来确定垂向力的参数 (只考虑了 $z$ 方向，忽略了 $R$ 方向对 $z$ 方向的关联)。即通过使用 $\chi^2$ 这种积分数据处理方法使建模的量和观测的量之间的差异最小化：

$$\chi^2 = \chi_\nu^2 + \chi_\sigma^2 + q(V_c(9.5\,{\rm kpc}) - V_c(7.5\,{\rm kpc}))^2 \ . \tag{114}$$

Bienaymé 等人[52]得到的太阳附近盘组分的面密度 $\Sigma_{\rm disc}(R_0) = (44.4 \pm 4.1)\,M_\odot \cdot {\rm pc}^{-2}$，太阳附近的暗物质密度 $\rho_{\rm DM}(z = 0) = (0.014\,3 \pm 0.001\,1)\,M_\odot \cdot {\rm pc}^{-3}$，Oort 极限下重子密度和总密度分别为 $\rho_{\rm baryons}(z = 0) = (0.077 \pm 0.007)\,M_\odot \cdot {\rm pc}^{-3}$ 和 $\rho_{\rm baryons}(z = 0) = (0.091 \pm 0.005\,6)\,M_\odot \cdot {\rm pc}^{-3}$。

此外，Bienaymé 等人[53]将一般轴对称引力势近似并拼凑为 Stäckel 势形式，即令 $V(R,z)=\tilde{V}(\lambda,\nu)$ 为在位置 $(R_1,z_1)=(\lambda_1,\nu_1)$ 处局部近似的真实引力势，其 Stčkel 势为 $\varPhi(\lambda,\nu)$ (详见其附录 A)。并以和前述同样的方法基于 Stäckel 势假定了 $I_3$ 的表达式：

$$I_3=\varPsi(R,z)-\frac{1}{2}\frac{L^2-L_z^2}{z_0^2}-\frac{1}{2}v_z^2\ , \tag{115}$$

其中，$\varPsi(R,z)=-\left[V(R,z)-V(\sqrt{\lambda},0)\right]\dfrac{\lambda+z_0^2}{z_0^2}$。

Bienaymé[54]还基于 Stäckel 势近似给出了一般三轴引力势下的估算第二、第三运动积分 $I_2$, $I_3$ 的一种近似的解析表达式 (见本文 2.5.1 小节)。然后讨论了基于此的分布函数模型：

$$f=\exp\left\{-n\left[E+\frac{1}{r_E^2}(bI_2+cI_3)\right]\right\}\ , \tag{116}$$

并用 Jeans 方程进行了检验。

3.2.2 一个更多参数的分布函数模型

下面介绍一种参数更多、复杂的分布函数模型[40, 42]。Famaey 等人[42]构建基于 3 个运动积分的分布函数 $f(E,\ L_z,\ I_3)$，从而实际星系的恒星盘通过这种形式的分布函数 (或者若干此种分布函数的线性组合) 来正确地拟合出。他们的方法是，(1) 对于 $I_3$ 的计算，简单套用 Stäckel 势的 $I_3$ 表达式；(2) 对于分布函数的形式，在前人的一个基于 2 个运动积分的分布函数[40]的基础上，引入更多自由参数作拓展。

分布函数的具体形式为：

$$F(E,L_z,I_3)=f(L_z)\left(\frac{E-S_{z_0}(L_z)}{S_z(L_z)-S_{z_0}(L_z)}\right)^{\eta}(p+qE+rL_z^2+sI_3)^{\delta}\ , \tag{117}$$

其中，$f(L_z)=\dfrac{1}{1+\mathrm{e}^{-aL_z}}(2L_z^2S_{z_0}(l_z))^{\beta}(S_{z_0}(L_z))_{\alpha_1}\mathrm{e}^{-\frac{\alpha_2}{S_{z_0}(L_z)}}$，$E$, $L_z$, $I_3$ 为 3 个运动积分，$p$, $q$, $r$, $s$ 为参数。根据这种分布函数，在 $(E,L_z^2)$ 平面，若保持 $z=z_0$ 固定，允许改变角频率，则得到边界线族 $E\geqslant\varPhi-\dfrac{L_z^2}{2\bar{\omega}_0^2}(L_z=v-\phi\bar{\omega})$；可以用 $E=S_{z_0}(L_z)$ 表示包络线。

通过计算分布函数下若干阶速度的平均，即可得到该分布函数 (运动积分空间) 的矩：

$$\mu_{l,m,n}(\lambda,\nu)=\int\int\int F(E,L_z,i_3)v_\lambda^l v_\phi^m v_\nu^n \mathrm{d}v_\lambda \mathrm{d}v_\mu \mathrm{d}v_\nu\ , \tag{118}$$

从而 0 阶矩为质量密度如 $\rho(\lambda,\nu)=\mu_{0,0,0}(\lambda,\nu)$，二阶矩为速度弥散如 $\rho\sigma_\nu^2(\lambda,\nu)=\mu_{0,0,2}(\lambda,\nu)$。文献 [42] 的第 4 章给出了一种方法，通过由积分区域边界的线性组合表示积分区域中的点，来计算其各阶矩中的积分；本文附录 B 简介了其基本思想。有了矩的计算，文献 [42] 的第 5 章介绍并分析了分布函数中各种参数所体现的物理特征。

文献 [42] 最后介绍了对实际的恒星盘建模的方法：通过将分布函数组分作为基础函数线性组合。这需要在配置空间中引入一个网格 $(\bar{\omega}_i,z_i)$，并最小化方差 (如对于分量 $F_\varLambda$ 代表

0 阶矩，即密度 $\mu^{(\Lambda)}_{0,0,0}(\bar{\omega}_i, z_i)$ 时)：

$$\chi^2 = \sum_i \left[ (\sum_\Lambda c_\Lambda \mu^{(\Lambda)}_{0,0,0}(\bar{\omega}_i, z_i)) - \rho_0(\bar{\omega}_i, z_i) \right]^2 , \tag{119}$$

其中，$\Lambda = (\alpha_1,\ \alpha_2,\ \beta,\ \delta,\ \eta,\ z_0,\ s,\ a)$ 为分布函数的参数集，$c_\Lambda$ 为系数。拟合程序中不断加入满足要求特征的新组分，直至 $\chi^2$ 的最小值不显著变化。

# 4 讨论和总结

前面内容聚焦于准稳态时星系的分布函数特点及建模。但是，与常见的稳态、准稳态的热力学系统相比，自引力系统有些特点本质不同，这些特点也许会限制前述方法的适用范围。另一方面，当考虑处于非稳态 (远离平衡时) 的星系结构和演化过程，或者考虑少体引力系统 (粒子数较少的引力系统) 中的具体恒星的运动，上述基于作用量的建模方法是否有效，或者说如何使得上述方法仍能应用于某些具体场合，这是非常值得探索的课题。

在本章里，我们将首先在动力系统理论的视角下 (dynamical-systems theory)，研究自引力系统的动力学 (见 4.1 节)；然后介绍和评述近年来研究者把“作用量-角变量”方法应用到完全不可积的星系现象和星系结构的进展 (共振和星系棒，见 4.2 节)，应用于非平衡的星系过程——潮汐星流 (tidal streams) 的研究，包括星流的形成过程，在作用量空间、角变量空间对星流的刻画，以及把星流用来拟合主星系的引力势分布 (见 4.3 节)；最后是全文的总结 (见 4.4 节)。

## 4.1 动力系统理论视角下的研究

从统计物理角度，引力 $N$ 体系统是一种长程相互作用系统。与一般的近程作用系统不同，这类系统中任何粒子之间的作用几乎都不能忽略 (粒子间的作用势 $\psi \propto r^{-\alpha}$，且 $\alpha$ 小于位形空间维数 $d$；假设粒子在全空间均匀分布，那么总作用势 $V \propto r^{d-\alpha}|^{\infty}_{r_0}$ 将是发散的)，并且与势能有关的宏观量不再具有统计物理所经常要求的可加性。另一方面，自引力系统由于可形成核-晕结构使其熵可能不具有最大值，因此并没有一个真正的平衡态；一般来说，相比于近程作用系统，自引力系统的弛豫过程可以更快，例如剧烈弛豫机制 (violent relaxation)[55, 58]。这些特点导致自引力系统与统计物理中常见的近程作用系统显著不同，往往不能用平衡态统计方法来很好地研究。

从动力学角度，作用量-角变量坐标将 $n$ 自由度可积系统的 $2n$ 维相空间变成 $n$ 维环面 $T^n$，如果存在一组非零整数 $k_1,\ k_2,\ k_3, \cdots ,\ k_n$，使得角频率 $\boldsymbol{\Omega}$ 满足 $\sum^n k_i \Omega_i = 0$ 时，则存在共振。共振关系可有不同多种。庞加莱 (Poincaré) 截面法就是在相空间 (即坐标维+速度维) 中取一个低于相空间维度的“面”，记录系统的轨道穿过这个面的点，并研究这些落点规律，该方法是降维研究运动的常见且直观的方法。以3自由度可积系统为例 (如在 Stäckel 势形式的星系中的任意一个恒星轨道)，该系统在相空间中呈现为一个三维环体 (3-torus)；若系统的其中两个频率比 $\Omega_1/\Omega_2$ 为有理数，则在其相应的二维环面的运动是周期的，其在

环面上的运动轨迹为周期轨线，在庞加莱截面上则是一些离散的不动点；若运动是准周期的，则是布满环面的准周期轨线，落点在截面上将形成一条闭合曲线。

当可积系统受不可积微扰且成为近可积系统 (近可积系统是大部分的相空间仍被规则轨道占据的受扰可积系统)，KAM 定理给出了近可积系统中仍然稳定存在局部的不变环面的条件。KAM 定理是 20 世纪中叶 Kolmogorov, Arnold 和 Moser 提出的重要定理，后续的工作放宽了定理的条件。对于截面上的不动点，当施加扰动，简单来说，经过多次运动后，双曲 (不稳定) 不动点将逐渐偏离，椭圆 (稳定) 不动点则将围绕不动点形成闭合环面，或在满足 KAM 定理的一定区域内 (即 KAM 环面) 稳定存在，或形成更高阶的新的两种不动点。如此迭代，将导致更复杂和不可预知的演化，从而产生混沌。此外，对于 $n > 2$ 自由度的近可积系统，相空间为 $2n$ 维，等能面 (或称能量流形) 为 $(2n-1)$ 维，而 $n$ 维的 KAM 环面不再能将等能面分割成闭区域的集合，即存在阿诺德扩散现象 (Arnold diffusion)。阿诺德扩散使得 $n \geqslant 3$ 自由度的动力系统的行为更为复杂，在稳定的轨道之间存在着高度不稳定的轨道，这样运动轨迹就可以绕过 KAM 环面到其他区域，仍会出现混沌；人们对于阿诺德扩散的理解，目前仍然非常有限。大致上，我们可以说，有理环面的破坏源于运动模式的非线性共振，非线性共振进一步导致混沌运动[63]。现代的“动力系统”学科本身也在拓扑和分析上有了更严密的定义和更多的理论，我们期待着这些现代理论能在自引力系统的动力学研究中大放异彩。

### 4.2 应用于“不可积”的星系现象

绝大多数自引力系统 (如星系) 都不是可积系统；即使数值计算，由于混沌的存在，也几乎不可能精确计算某个具体天体 (如某个恒星) 实际的长期运动。尽管如此，对于处于准稳态的恒星集体 (即星系) 而言，前述章节所介绍的 Binney 等人的工作表明：恒星集体的统计性质，例如各种分布，是可以利用“作用量-角变量”方法来可靠地获得的。我们进一步感兴趣的是，“作用量-角变量”方法是否可以应用于共振、混沌等所谓“不可积的”的恒星运动，可以应用于哪些具体场合？最近 Binney[16, 17]研究了把环面映射方法应用到棒旋星系的共振轨道的情况 (以银河系观测数据为实例)。他在文献 [16] 中，分别研究了在银河系棒的这种扰动强度下，“作用量-角变量”工具对于外林德布拉德共振、共转共振、内林德布拉德共振的适用性。他发现，对于银河系被外林德布拉德共振和共转共振所俘获的恒星轨道，环面映射法能很好地拟合出；但是更靠近棒内部的内林德布拉德共振俘获的轨道，环面映射法就显得不够好了。Binney[17]更进一步把上述方法应用到太阳邻域恒星的速度空间数据中，发现在银河系的棒模型下，环面映射法所拟合的共转共振情形与观测数据一致，而外林德布拉德共振情形不被数据所支持；也就是说，基于环面映射法和微扰理论的计算，对于棒旋星系中的共振俘获现象，可以有很好的实用性了。

### 4.3 应用于远离平衡的星系过程：潮汐星流

当一个星团或矮星系 (称为前身天体) 靠近一个大星系 (称为主星系) 时，前身天体各个不同位置所受的引力不同，从而发生形状和速度的变化；当所处的引力场梯度足够大时，前身天体的恒星们会被主星系巨大潮汐力逐步拉扯，或多或少的恒星脱离前身天体，常形成长

条状或称纤维状的结构，此即潮汐星流 (详见文献 [33])。

研究者发现，对于星流的这种形成过程，以及对于星流特性的刻画，最方便的方式是基于“作用量-角变量”体系[34, 35]。当前身矮星系或星团的成员恒星没有被潮汐剥离，随着前身天体在主星系引力势中运动时，在主星系参考系中，这些恒星的作用量是不守恒的。但一旦某恒星被潮汐力剥离前身天体之后，就可以忽略前身天体以及星流中的其他恒星对它的引力作用，认为只是在主星系的引力势下运动。那么，自剥离时刻起，该恒星的作用量就一直保持剥离前的数值，不再随时间变化了。下面我们沿用 Eyre 和 Binney[36] 的记号和表述方式，来简介这个方法。设前身天体 (小星系或星团) 位于固定背景势下的规则轨道上，其作用量为 $\boldsymbol{J}_0$。背景势具有哈密顿量 $H(\boldsymbol{J})$，其中 $\boldsymbol{J}$ 是潮汐星流中的某恒星的作用量，一般与 $\boldsymbol{J}_0$ 相差不远。对于潮汐星流中的某个恒星，其作用量是常量，角变量随时间线性增加。在 $\boldsymbol{J}_0$ 附近，对哈密顿量作泰勒展开：

$$H = H_0 + \boldsymbol{\Omega} \cdot \delta \boldsymbol{J} + \frac{1}{2} \delta \boldsymbol{J}^{\mathrm{T}} \boldsymbol{D} \cdot \delta \boldsymbol{J} \ , \tag{120}$$

其中 $\delta \boldsymbol{J} = J - J_0$；$\boldsymbol{\Omega}_0$ 是前身天体的轨道的频率，$\Omega_{0,i} = \frac{\partial H}{\partial J_i}|_{\boldsymbol{J}_0}$；$\boldsymbol{D}$ 是哈密顿量的海森矩阵 (即 $H$ 的二阶微分)，$D_{ij} = \dfrac{\partial^2 H}{\partial J_i \partial J_j}|_{\boldsymbol{J}_0}$。则位于前身天体轨道附近的某恒星的轨道频率为：

$$\boldsymbol{\Omega}(\boldsymbol{J}) = \boldsymbol{\Omega}_0 + \boldsymbol{D} \cdot \delta \boldsymbol{J} \ . \tag{121}$$

关于星流的成员恒星，它们的作用量和角变量都是有着某种分布，即某种程度的弥散，$\Delta \boldsymbol{J}$ 和 $\Delta \boldsymbol{\theta}_0$。由前述可知，成员恒星的作用量弥散 $\Delta \boldsymbol{J}$ 自从星流形成后就不变了；但角变量弥散不会保持在初始的 $\Delta \boldsymbol{\theta}_0$，而是将随时间线性增大。角变量弥散的关系式为[34, 36]：

$$\Delta \boldsymbol{\theta}(t) = t \Delta \boldsymbol{\Omega} + \Delta \boldsymbol{\theta}_0 \approx t \Delta \boldsymbol{\Omega} \ , \tag{122}$$

其中约等号是由于潮汐流形成后，$\boldsymbol{\theta}_0$ 将远小于正比于时间的那一项。于是有

$$\Delta \boldsymbol{\Omega}(\boldsymbol{J}) = \boldsymbol{D} \cdot \Delta \boldsymbol{J} \ . \tag{123}$$

上式表明海森矩阵 $\boldsymbol{D}$ 的作用是：把恒星们在作用量空间的分布 (参见文献 [36] 的图 10) 线性映射到角变量空间的分布 (参见文献 [36] 的图 13)。$\boldsymbol{D}$ 是一个对称矩阵，因此其特征取决于其相互正交的特征向量 $\hat{\boldsymbol{e}}_n$ $(n = 1,\ 2,\ 3)$，和实特征值 $\lambda_n$[36]。海森矩阵决定了星流的拉伸方向，即星流沿着最大特征方向拉长[20]：

$$\frac{\Delta \boldsymbol{\theta}}{t} \approx \Delta \boldsymbol{\Omega} \approx \hat{\boldsymbol{e}}_1 (\lambda_1 \hat{\boldsymbol{e}}_1 \cdot \Delta \boldsymbol{J}) \ . \tag{124}$$

什么样的引力势分布形式 (必要条件) 可以形成星流等技术细节，这里不作详细介绍，读者可以参考 Sanders 的博士论文 (第 5 章和第 6 章)[20]。特别是该论文第 6 章详细介绍了如何应用星流数据来拟合限定主星系的引力势分布。

最后，我们要指出的是：当前身天体为很小的星系或者星团时，星流的轨迹可以粗略地 (1 阶近似) 看作是前身天体在主星系中的轨道；因此，过去研究者常简单假定星流为一条轨

道，从而拟合主星系引力势，此即文献中所说的轨道拟合法 (例如文献 [38])。但是，基于上面的分析，星流的轨迹并不必然与前身天体的轨道或者任何其他轨道重合 (参见文献 [36] 的图 13)；因此轨道拟合法会导致较大的主星系引力势偏差 (详见文献 [20] 的第 6 章)。

### 4.4 结束语

本文主要目的是介绍星系动力学建模方面的新进展，即作用量计算方法，以及近年来基于作用量 (或运动积分) 的分布函数建模在银河系数据上的应用，另外还简介了“作用量-角变量”方法在不可积的星系现象 (共振和星系棒) 和非平衡过程 (例如潮汐星流的形成) 的应用，以期为中国在“Gaia＋LAMOST”黄金时代的银河系研究做些贡献。目的之二，厘清上述方法及其应用的理论基础和脉络。在实际或理想的星系中，唯一的一类可积系统是具备 Stäckel 形式引力势的星系；相应地，几乎所有星系建模工作的解析性基础，都是 Stäckel 势理论。因此，我们还在本文第 1 章花不少篇幅从 $N$ 个粒子的经典力学 (哈密顿方程和刘维尔方程) 出发，依次引入平均场近似，介绍无碰撞玻尔兹曼方程；然后在经典力学的正则变换以及相空间的语境下，引入运动积分、正则环面坐标即作用量和角变量；接着就介绍唯一的严格可积的 (理想的)“星系”，即介绍 Stäckel 势理论的各种解析式。介绍这些来龙去脉，期望我国学者不仅能应用 Binney 等国外专家发明发展的方法来分析银河系数据，还能够在理论基础和方法上同样有所创新。

### 参考文献：


[1] de Zeeuw P T. MNRAS, 1985, 216: 273
[2] de Zeeuw P T, Lynden-Bell D. MNRAS, 1985, 215: 713
[3] Lynden-Bell D. MNRAS, 1962, 124: 95
[4] Dejonghe H, de Zeeuw P T. ApJ, 1988, 329: 720
[5] Dejonghe H, de Zeeuw P T. ApJ, 1988, 333: 90
[6] Dejonghe H. ApJ, 1989, 343: 113
[7] Dehnen W, Binney J. MNRAS, 1998, 294: 429
[8] Binney J, Tremaine S. Galactic Dynamics. 2nd ed. Princeton, New Jersey: Princeton University Press. 2008
[9] Stäckel P. Math. Ann., 1890, 35: 91
[10] Stäckel P. Über die Integration der Hamilton-Jacobischen Differential Gleichung mittelst Separation der Variabein, Habilitationsschrift: Halle. 1891
[11] McGill C, Binney J. MNRAS, 1990, 244: 634
[12] Binney J, McMillan P J. MNRAS, 2016, 456: 1982: 1998
[13] Binney J. MNRAS, 2012, 426: 1324
[14] Binney J. MNRAS, 2012, 426: 1328
[15] Binney, J. IAU, 2020, 353: 101
[16] Binney J. MNRAS, 2020, 495: 886
[17] Binney J. MNRAS, 2020, 495: 895
[18] Binney J, MicMillan P J. TM: Torus Mapper. ascl, 2015, 1512.014
[19] Sanders J L. MNRAS, 2012, 426: 128
[20] Sanders J L. Dynamics of the Milky Way: Tidal Streams and Extended Distribution Functions for the

Galactic Disc. Switzerland: Springer International Publishing, 2015
[21] Sanders J L, Binney J. MNRAS, 2014, 441: 3284
[22] Sanders J L, Binney J. MNRAS, 2015, 447: 2479
[23] Sanders J L, Binney J. TACT: The Action Computation Tool. ascl, 2015, 1512.020
[24] Sanders J L, Binney J. MNRAS, 2016, 457: 2107
[25] van de Ven G, Hunter C, Verolme E K, et al. MNRAS, 2003, 1056: 1082
[26] Pontzen A, Governato F. MNRAS, 2013, 430: 121
[27] Posti L, Binney J, Nipoti C, et al. MNRAS, 2015, 447: 3060
[28] Sellwood J A, Binney J. MNRAS, 2002, 336: 785
[29] Bovy J. ApJ, 2014, 795: 95
[30] Bovy J. galpy: Galactic dynamics package. ascl, 2014, 1411.008
[31] Vasiliev E. MNRAS, 2019, 482: 1525
[32] Ibata R, Diakogiannis F I, Famaey B, et al. Arxiv, 2020, 2012: 05250
[33] Helmi A. ARAA, 2020, 58: 205
[34] Helmi A, White S D M. MNRAS, 1999, 307: 495
[35] Tremaine S. MNRAS, 1999, 307: 877
[36] Eyre A, Binney J. MNRAS, 2011, 413: 1852
[37] Sanders J L, Binney J. MNRAS, 2013, 433: 1813
[38] Binney J. MNRAS, 2008, 386: L47
[39] Kuijken K, Gilmore G. MNRAS, 1989, 239: 571
[40] Batsleer P, Dejonghe H. A&A, 1995, 294: 693
[41] Ivezié Z, Sesar B, Jurić M, et al. ApJ, 2008, 684: 287
[42] Famaey B, Van Caelenberg K, Dejonghe H. MNRAS, 2002, 335: 201
[43] Bovy J, Allende Prieto C, Beers T C, et al. ApJ, 2012, 759: 131
[44] Bovy J, Rix H-W. ApJ, 2013, 779: 115
[45] Binney J. MNRAS, 2010, 401: 2318
[46] Binney J, McMillan P. MNRAS, 2011, 413: 1889
[47] Binney J, Burnett B, Kordopatis G, et al. MNRAS, 2014, 439: 1231
[48] Binney J. MNRAS, 2014, 440: 787
[49] Binney J, Piffl T. MNRAS, 2015, 454: 3653
[50] Piffl T, Binney J, McMillan P J, et al. MNRAS, 2014, 445: 3133
[51] Piffl T, Penoyre Z, Binney J. MNRAS, 2015, 451: 639
[52] Bienaymé O, Famaey B, Siebert A, et al. A&A, 2014, 571: A92
[53] Bienaymé O, Robin A C, Famaey B, et al. A&A, 2015, 581: A123
[54] Bienaymé O. A&A, 2019, 627: A123
[55] Lynden-Bell D. MNRAS, 1967, 136: 101
[56] Lynden-Bell D, Wood R. MNRAS, 1967, 138: 495
[57] McMillan P J. MNRAS, 2017, 465: 76
[58] Padmanabhan T. PhR, 1990, 188: 285
[59] Levin Y, Pakter R, Rizzato F B, et al. PhR, 2014, 535: 1
[60] Pottier. Nonequilibrium Statistical Physics - Linear Irreversible Processes. Oxford: Oxford University Press. 2010: 89
[61] Read J I. Journal of Physics G Nuclear Physics, 2014, 41: 063101
[62] Arnold V I. 经典力学的数学方法 (齐民友译). 北京: 高等教育出版社, 2005: 125
[63] 郑志刚, 胡刚. 从动力学到统计物理学. 北京: 北京大学出版社, 2016: 15
[64] Florin Diacu, Philip Holmes. Celestial Encounters: The Origins of Chaos and Stability. Princeton: Princeton University Press. 1996: 193
[65] Sinai Y G. Lectures in ergodic theory. Lecture notes in Mathematics. Princeton: Princeton University Press, 1977, 4: 1121

[66] Beraldo e Silva L, de Siqueira Pedra W, Valluri M, et al. ApJ, 2019, 870: 128
[67] Beraldo e Silva L, de Siqueira Pedra W, Valluri M. ApJ, 2019, 872: 20
[68] Lange S, Richter M, Onken F, et al. CHAOS, 2014, 24: 024409
[69] Lilley E J, Evans N W, Sanders J L. MNRAS, 2018, 476: 2086
[70] Power C, Knebe A, Knollmann S R. MNRAS, 2012, 419: 1576
[71] Klypin A, Yepes G, Gottlöber S, et al. MNRAS, 2016, 457: 4340
[72] Bullock J S, Boylan-Kolchin M. ARA&A, 2017, 55: 343

# Extensions of Stäckel Gravitational Potential Theory and Their Application in Galactic Modeling

GU Jian-yu[1,2], DONG Xiao-bo[1,2]

*(1. Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China)*

**Abstract:** Stäckel gravitational potentials are the most general class of separable potentials. Galaxies with a Stäckel form belong to complete integrable systems, in which the stellar orbits are all regular and their integrals of motion ($\boldsymbol{I}$) can be expressed analytically. Integrals of motion, particularly actions ($\boldsymbol{J}$) as a special kind of integrals, can simplify the description of stellar motions. An important approach to study the kinematics and dynamics of stars in galaxies is by means of action space; e.g., using actions as the arguments of distribution functions (DFs), $f(J_R, L_z, J_z)$. One can estimate actions for the stellar motions in general galaxies by locally approximating the gravitational potentials with Stäckel models. Recent years, there have been important progresses in extending the use of Stäckel potential theory in general galaxies: in the respect of computing actions (or integrals of motion) for general galaxies, researchers have developed several fast algorithms (e.g., Stäckel Fudge and Stäckel Fitting), and several convergent high-precision algorithms based on torus mapping (e.g., generating function from orbit integration method and iterative torus construction method), as well as formalisms of analytically expressing “approximate integrals of motion” of general galaxies (e.g., the quasi-$I_3$); in the respect of constructing DFs for galaxies and galactic components, with the actions estimated by the above methods or directly using integrals formulae of Stäckel potentials (particularly the $I_3$ expression), some DF models are proposed in the form of $f(J_R, L_z, J_z)$ or $f(E, L_z, I_3)$. These theoretical and methodological developments make possible DF modelling for galaxies. This article reviews the above developments, and introduces several recent representative studies that use action- or integral-based DF models to fit the MW data, as well as the recent developments in extending the use of the (action,

angle) methodology to non-integrable phenomena (such as galactic resonances and bars) and out-of-equilibrium processes (such as the formation of tidal streams), with a hope of being helpful for Chinese astronomical community to welcome this “Gaia + LAMOST” era.

**Key words:** Stäckel gravitational potentials; action-angle variables estimation; Milky Way distribution function modeling

## 附录 A 计算作用量、角频率和角变量的步骤

这里介绍 Sanders 和 Binney 的程序中数值计算作用量、角频率和角变量的步骤 (文献 [1] 中第 8 节及附录；文献 [22] 的附录 A)。程序中的这些中间量的计算以及轨道分类，来自 Sanders 和 Binney 的 TACT 程序[23] 的 /tact/aa/inc/stackel_aa.h 中的函数定义。

(1) 作用量 $J$

作用量的计算公式为

$$J_\tau = \frac{1}{2\pi}\oint \frac{\partial W_\tau}{\partial \tau}\mathrm{d}\tau = \frac{1}{2\pi}\oint p_\tau \mathrm{d}\tau = \frac{2}{\pi}\int_{\tau_-}^{\tau_+} |p_\tau|\mathrm{d}\tau \ , \quad \tau = \lambda,\ \mu,\ \nu, \tag{125}$$

其中，$\tau_-$, $\tau_+$ 是 $p_\tau^2 = 0$ 时的根。注意对于有些环轨道，$\tau$ 的积分上下限可能是 $-\alpha$, $-\beta$, $-\gamma$，具体可见文献 [22]的表 1，其详细讨论可见文献 [1] 中的轨道分类内容。

(2) 角频率 $\varOmega$

由哈密顿方程 ($\varOmega_\tau = \dot{\theta} = \partial H/\partial J_\tau$) 以及微分关系，可以得到[1]：

$$\begin{cases} \dfrac{\partial E}{\partial E} = 1 = \varOmega_\lambda \dfrac{\partial J_\lambda}{\partial E} + \varOmega_\mu \dfrac{\partial J_\mu}{\partial E} + \varOmega_\nu \dfrac{\partial J_\nu}{\partial E} \\ \dfrac{\partial E}{\partial I_2} = 0 = \varOmega_\lambda \dfrac{\partial J_\lambda}{\partial I_2} + \varOmega_\mu \dfrac{\partial J_\mu}{\partial I_2} + \varOmega_\nu \dfrac{\partial J_\nu}{\partial I_2} \\ \dfrac{\partial E}{\partial I_3} = 0 = \varOmega_\lambda \dfrac{\partial J_\lambda}{\partial I_3} + \varOmega_\mu \dfrac{\partial J_\mu}{\partial I_3} + \varOmega_\nu \dfrac{\partial J_\nu}{\partial I_3} \end{cases} . \tag{126}$$

将这个关于角频率的线性方程反解可得角频率，如

$$\varOmega_\lambda = \frac{1}{\varGamma}\frac{\partial(J_\mu, J_\nu)}{\partial(I_2, I_3)} \ , \qquad \varGamma = \frac{\partial(J_\lambda, J_\mu, J_\nu)}{\partial(E, I_2, I_3)} \ , \tag{127}$$

可用式 (125) 与动量和运动积分的关系计算：

$$\frac{\partial J_\tau}{\partial I} = \frac{1}{2\pi}\oint \frac{\partial p_\tau}{\partial I}\mathrm{d}\tau \ , \quad I = E,\ I_2,\ I_3 \ . \tag{128}$$

其中，由 Stäckel 势中动量和运动积分的关系可得：

$$\begin{cases} \dfrac{\partial p_\tau}{\partial E}|_\tau = \dfrac{\tau^2}{4p_\tau(\tau+\alpha)(\tau+\beta)(\tau+\gamma)} \\ \dfrac{\partial p_\tau}{\partial j}|_\tau = -\dfrac{1}{\tau}\dfrac{\partial p_\tau}{\partial E}|_\tau \\ \dfrac{\partial p_\tau}{\partial k}|_\tau = \dfrac{1}{\tau^2}\dfrac{\partial p_\tau}{\partial E}|_\tau \end{cases} . \tag{129}$$

注意上式右端作为分母的 $p_\tau$，在积分边界会消失。Sanders 和 Binney[22] 作了下述变换：

$$\tau = \hat{\tau}\sin\theta + \bar{\tau}\ ,\ \ \bar{\tau} = \frac{1}{2}(\tau_- + \tau_+)\ ,\ \ \hat{\tau} = \frac{1}{2}(\tau_- - \tau_+)\ , \tag{130}$$

使被积函数在积分边界更平滑地变到零。这样，每一点积分变量的值都得到了，再数值积分得到角频率。程序中用的是高斯-勒让德正交格式来进行积分。

(3) 角变量 $\theta$

用哈密顿特征函数计算角变量，表示为：

$$\theta_\tau = \frac{\partial W}{\partial J_\tau} = \sum_{I=E,I_2,I_3}\frac{\partial W}{\partial I}\frac{\partial I}{\partial J_\tau}\ . \tag{131}$$

计算 $\partial W/\partial I$ 所用的 $W(\boldsymbol{\tau},\boldsymbol{J})$ 为：

$$W = \sum_\tau W_\tau = \sum_\tau \int_{\tau_-}^{\tau} \mathrm{d}\tau' p'_\tau + \mathscr{F}_\tau(p_\tau,\boldsymbol{x})\int_{\tau_-}^{\tau_+}\mathrm{d}\tau'|p'_\tau|\ , \tag{132}$$

其中，被积函数的处理类似于前述所示；函数 $\mathscr{F}_\tau(p_\tau,\boldsymbol{x})$ 是考虑轨道类型的特点，为消除椭球坐标中的简并性以使角变量覆盖笛卡尔坐标中一个振动的全部范围而引入的 (文献 [22] 附录中的式 (A6))，即计算中还要进行轨道方面的判断和修正。

# 附录 B　Famaey 等人[42]的运动积分分布函数模型的矩的计算处理方法简介

对于式 (118)，从被积的速度空间转换到 3 个运动积分所构成的空间

$$\begin{aligned}\mu_{l,m,n} = &\frac{2^{\frac{l+n}{2}+1}}{\omega(\lambda-\nu)^{\frac{l+n}{2}}}\int\frac{f(L_z^2)}{\sqrt{L_z^2}}v_\phi^m\mathrm{d}L_z^2\int\int\mathrm{d}E\mathrm{d}I_3\int\int\left(\frac{E-S_{z0}}{S_0-S_{z0}}\right)^\eta \\ &(p+qE+rL_z^2+sI_3)^\delta(I_3^+-I_3)^{\frac{l-1}{2}}(I_3-I_3^-)^{\frac{n-1}{2}}\ ,\end{aligned} \tag{133}$$

其中，$I_3^+(E,L_z^2) = (\lambda+\gamma)[G(\lambda)-E] - \dfrac{\lambda+\gamma}{2(\lambda+\alpha)}L_z^2$，$I_3^-(E,L_z^2) = (\nu+\gamma)[G(\nu)-E] -$

$\dfrac{\nu+\gamma}{2(\nu+\alpha)}L_z^2$。在 $(E, I_3)$ 面的积分区域为：

$$\begin{cases} I_3^-(E, L_z^2) \leqslant I_3 \leqslant I_3^+(E, L_z^2) \\ S_{z_0}(L_z) \leqslant E \leqslant \phi - \dfrac{L_z^2}{2\omega^2} \\ p + qE + rL_z^2 + sI_3 \geqslant 0 \end{cases} . \tag{134}$$

通过对积分边界线 (上式取等号时) 的线性组合，来表示式 (118) 中的积分区域中的点，

$$\frac{E - S_{z_0}}{S_0 - S_{z_0}} = t(I_3^+ - I_3) + u(I_3 - I_3^-) + v(p + qE + rL_z^2 + sI_3) \ . \tag{135}$$

之后，将 $(E,\ I_3)$ 的积分区间转换到新的积分变量空间 (固定 $L_z$)，然后对比变换前后的每一项，解出上式中的组合系数 $v$, $u$, $t$ 的关系，从而用一维数值积分计算其矩 (更多细节见文献 [42] 中的 4.2 和 4.3 节)。